Inequality at the Top: Down to the Roots

Riccardo A. Cioffi, George L. Sorg-Langhans, Maximilian J. Vogler

November 20, 2020

Abstract

Multiple theories of inequality compete to explain U.S. wealth inequality and the share of wealth held by the top one percent. To what extent does it matter which of these models we rely on? In this paper we analyze the responses of the different theories to a host of policy experiments. To this end, we form a quantitative model that nests the competing channels and assesses the effects of policy experiments by sequentially shutting off all but one of these model mechanisms. Our model is calibrated on the wealth distribution which allows us to starkly contrast the different theories and clearly understand the mechanisms at work. Our policy exercises allow us to compare the competing models against empirical estimates from Jakobsen *et al.* (2020) and lead us to conclude that heterogeneous returns is the most important cause of increased wealth inequality. Understanding this dominant channel is crucial for both policy considerations and academic modeling.

1 Introduction

Multiple theories of inequality compete to explain U.S. wealth inequality and the share of wealth held by the top one percent. The purpose of this paper is to tease these channels apart through a quantitative investigation by assessing their response to a sequence of policy experiments and shocks. To this end we build a quantitative model that nests many of these competing theories, calibrate it, and sequentially shut down subsets of the channels. The resulting restricted models' responses to the various policy experiments then allow us to discriminate between different theories.[1] We find that the key variables of interest, inequality and top wealth share, behave significantly different in the context of the competing channels.

The main contribution of our paper lies in teasing the different channels apart based on their predictions in various policy experiments. We then contrast these predictions with empirical results from recent research by Jakobsen et al. (2020). This allows us to

[1] We will denote as "complete model" the model in which all channels are present, and as "restricted models" the models in which only a subset of the channels is present.

identify the dominant channel of wealth inequality, which aids both policy makers and researchers.

We consider four distinct channels that give rise to high wealth concentration at the top. First, since business and non-financial assets represent 49% of portfolio holdings of the top one percent, our model features an entrepreneurship channel. In particular, we follow the mechanism proposed by Cagetti and De Nardi (2006): entrepreneurs face endogenous borrowing constraints due to limited commitment, which implies that they will rationally save more both to loosen the constraint and to grow their businesses.

Secondly, we consider the bequest channel, as proposed by De Nardi (2004). The relevance of this channel is highlighted in the work of De Nardi, French *et al.* (2010) and Dynan *et al.* (2004): Dynan *et al.* (2004) find a positive correlation between total lifetime income and savings rates, while De Nardi, French *et al.* (2010) show that the wealthy run down their assets at a slower rate than the wealth poor, leaving large estates to their offspring. In our model the bequest channel is implemented by the inclusion of a warm glow bequest motive. Here, leaving an estate to one's offspring is a luxury good, causing wealthier households to exhibit higher savings rates than their poor counter parts.

The third channel we nest in our model is heterogeneity in agents' discount factors. Krusell and A. A. Smith (1998) show that it is possible to significantly increase wealth inequality in a standard environment by including preference heterogeneity. The underlying mechanism is that households with lower discount factors tend to have higher consumption rates, which drives impatient households further into the left tail while patient households keep accumulating assets.

The last channel we consider is idiosyncratic income risk. As Castañeda *et al.* (2003) highlight, income inequality and risk are a significant contributing force to wealth inequality. Households exhibit high savings rates and hours worked in the high income states. This provides them with insurance against future low income realizations and drives inequality upwards. Policy experiments that dampen income inequality across states therefore lower the precautionary savings motive when households are in high income states. This provides an effective margin to distinguish this model from the other restricted models.

Since our aim is to arrive at clear and sharp predictions in the context of our channels we calibrate them directly on the wealth distribution. We consider matching the top wealth inequality as especially important because this is a key moment in the data. This calibration strategy will stretch the models' parameters and allow us to clearly understand the mechanisms at work and to contrast the predictions of the model.[2]

We consider two policy experiments in this paper. First, we consider a change in

[2]Even though these calibrations might not be entirely realistic, they do allow us to draw stark contrasts between the different models and provide a sharper answer to our motivating question.

capital gains taxation. This changes the effective rate of return on wealth, which decreases household incentives to save throughout the wealth distribution. Importantly, we find that indeed the models differ sharply in their prediction of how inequality moves in this scenario. Second, we investigate the effects of changing pensions in the various channels. This forces poor households to save for retirement, which decreases inequality. Interestingly this effect is dampened in the entrepreneurship model while it is catalyzed in the bequest model. Hence, we find different predictions for inequality across the model. Note that the counterpart to this policy experiment, an increase in the pensions rate does not have the same extreme effect. This is intuitive, since low consumption risk in the presence of Inada conditions has strong effects while an increase in consumption in the high states does not provide a similar effect.

The results from our first policy experiment allow us to uniquely identify the dominant cause of wealth inequality. Specifically, we observe that heterogeneous returns, as modeled by heterogeneous discount factors, yield different predictions for a change in the capital gains tax than any other channel. Due to the equivalence of capital gains and wealth taxes, this allows us to contrast these findings with the results of the empirical study in Jakobsen et al. (2020). In particular, these authors study the response of wealth inequality to a decrease in the wealth tax in Denmark and find that this policy increased wealth inequality. In our model, this pattern can only be matched by heterogeneous discount factors, which lends strong evidence to conclude that heterogeneous returns are the dominant driving force behind wealth inequality.

This result has important implications for both policy makers and economic researchers. First, for policy makers concerned about the increase in wealth inequality, knowing the underlying source enables a tailored policy response. Furthermore, it allows to understand the implicit tradeoffs inherent to such interventions. Second, this paper provides a guide for how to model wealth inequality for future economic research. Economists can thus reduce model complexity by only introducing heterogeneous discount factors to account for wealth inequality. In particular, this is of great value when economists study questions that are related to but not focused on wealth inequality. In addition, having identified the central channel should help focus future inequality research. For example, given the strength of this result, it is crucial to fully understand the origins of heterogeneous rates of returns.

The rest of this paper is organized as follows. Section 2 gives an overview of the literature. Section 3 outlines the nested model that combines all four channels and discusses a series of modeling choices and possible extensions. Our calibration strategy is documented in section 4. This section also contains our results for the replications of the core papers. Next, we discuss the results of the policy experiments in section 5. Finally, section 6 concludes.

2 Literature Review

The existence, nature, and consequences of shocks to household's earnings are of great significance for macroeconomic theory. After the work of Aiyagari (1994) and Huggett (1996), almost all models that try to account for the observed heterogeneity of wealth, include some form of earnings risk. A good overview of the different theories behind earnings risk and of their contribution to generating wealth inequality is Benhabib and Bisin (2017).

Once the issue of existence and importance of earnings risk had been settled, the literature moved in the direction of trying to understand what is its nature. After early empirical work, most quantitative macroeconomic models have focused on linear models of earnings dynamics in a Gaussian framework. Furthermore, these models often assume near random-walk behavior of the transitory component of income due to the simplicity of this formulation. In this tradition Castañeda *et al.* (2003) tries to explain wealth inequality by exploiting a high degree of earnings risk and is able to replicate the degree of wealth inequality observed in U.S. data. However, this is achieved at the cost of having a labor income process that is at odds with the data.

Recent work by Arellano, Blundell *et al.* (2017), Guvenen (2009), and Guvenen, Karahan *et al.* (2016) has cast doubts on the reliability of the linearity and Gaussianity assumptions . In particular Arellano, Blundell *et al.* (2017), using data from the PSID, find evidence of nonlinear persistence in earnings; Guvenen, Karahan *et al.* (2016), using data from the U.S. Social Security Administration, find evidence of non-Gaussianity; and Guvenen (2009) lends more support to what he defines as the Heterogeneous Income Profiles view, namely "individuals are subject to shocks with modest persistence, while facing life-cycle profiles that are individual-specific."

Let us now turn to the literature on entrepreneurship. The central reference for this paper is Cagetti and De Nardi (2006) as we adopt their approach for modeling entrepreneurship. In their model the optimal firm size is large enough to entice even wealthy entrepreneurs to keep saving to alleviate their borrowing constraints. We discuss the mechanism in more detail in the context of the policy experiments in section 5. Note that their model implements a less potent bequest motive than our model. To see this, note that the altruism model they use only generates a small wealth concentration when entrepreneurship is switched off.

Cagetti and De Nardi (2009) builds on their model and includes labor choice as a factor of production in the entrepreneur's problem. Note that the paper also considers the policy experiment of changing the bequest tax. The authors find that such a tax dampens the investment behavior of wealthy entrepreneurs leading to a decrease in inequality and a decrease in output. We use this finding in our design of the bequest tax experiment.

A key paper in the context of entrepreneurship is Quadrini (1999), as it represents a key empirical contribution by analyzing the link between entrepreneurship and high savings rates. Quadrini (2000) complements this empirical investigation by modeling entrepreneurship in a Bewley environment. Aside from generating a high share of capital holdings among the top one percent in the hands of entrepreneurs, the author is able to generate high social mobility for entrepreneurs, which is also found in the data.

One possible extension to our basic model is the inclusion of default in equilibrium for entrepreneurs. These bankruptcy considerations have been modeled by Glover and Short (2015). They use this environment to specifically consider changes to incorporation and the personal bankruptcy exemption. As these are not immediate policy experiments under consideration for this research question we will not explicitly model default risk in equilibrium.

In the following, we examine previous work regarding preference heterogeneity. There is a plethora of empirical evidence that supports the notion of heterogeneity in time preferences (e.g., see Cagetti 2003; Lawrence 1991).

The first model to relate preference heterogeneity to wealth inequality is the seminal paper by Krusell and A. A. Smith (1998). They enrich their infinite-horizon model with a stochastic process for discount rates. Even a small variation in time preferences is sufficient to match the variance of the wealth distribution and significantly improve the match for the top 1%, although they fall short of completely matching it. Krueger *et al.* (2017) follow the same tradition when using preference heterogeneity in order to match the wealth distribution, whose reaction to an aggregate shock they then proceed to examine.

In contrast to this approach, Hendricks (2007) studies preference heterogeneity in a life-cycle model with accidental bequests. Despite a higher variance of the discount factors, he finds that there is no significant improvement in the ability to match the top wealth shares. He argues that this divergence is due to the existence of an additional retirement savings motive in a life-cycle model. Paz-Pardo (2016) confirms Hendricks's findings but shows that a combination with the voluntary bequest channel of De Nardi (2004) manages to match the top wealth shares.

Finally, we survey the literature on the relevance of the bequest motive for wealth inequality. A range of papers has empirically examined the importance of the intergenerational transmission of wealth. Kotlikoff and Summers (1981), as well as Gale and Scholz (1994), argue that it is responsible for the majority of total wealth accumulation.

Given these facts, a theoretical literature has emerged that has furthered the understanding of the relation between bequests and wealth inequality. One important contribution is Becker and Tomes (1979), who model the bequest decision made by parents and conclude that parents, before leaving monetary bequests, first invest into a child's human capital.

De Nardi (2004) analyzes this problem in a quantitative framework. Specifically, she studies a multi-period OLG model that includes both human capital and monetary bequests. She also studies the importance of including a voluntary bequest motive through "warm glow" preferences in addition to accidental bequests. One implication of this structure is that bequests are a luxury good, so that the model succeeds at explaining the distribution of estates as only wealthy households leave bequests.

In contrast to these results, Hendricks (2004) and Hurd (1989) do not find evidence for a strong bequest motive using only accidental bequests. Furthermore, their choice of a utility function that is homothetic in bequests implies that not only rich but also poor people leave substantial bequests, which is at odds with the data. De Nardi and Yang (2014) show that, once you allow for voluntary bequests and the transmission of earnings ability, their model successfully matches the data.

A different approach is taken by Gokhale *et al.* (2001), who only use accidental bequests but enrich their model with a range of additional intergenerational links such as fertility, skill differences, assortative mating, social security and preference heterogeneity.

To our knowledge, there has so far not been extensive work on differentiating the channels of generating wealth inequality at the top. The paper that is closest to ours in terms of the research question is Benhabib, Bisin, and Luo (2016). They quantitatively examine which channels are necessary to reproduce the cross-sectional wealth distribution as well as social mobility in the U.S. The authors conclude that one needs a skewed distributions of earnings, differential savings and bequest rates across wealth levels, and capital income risk. In contrast to this paper, we do not ask which channels are theoretically necessary to replicate the wealth distribution. Rather, we examine a number of various channels and explore their quantitative implications to see whether they are consistent with the data.

3 Model

3.a *Environment*

Demographics

To best investigate the intergenerational links crucial to our project, we use a life-cycle version of the infinite horizon model. In order to keep the model computationally tractable, we adopt the perpetual youth framework developed by Blanchard (1985) and Gertler (1999).

Households can be young, middle-aged or old. Young households age with probability $1 - \lambda_y$ and middle-aged households age with probability $1 - \lambda_m$, whereas old households die with probability $1-\lambda_o$. There is a continuum of households of measure one facing idiosyncratic shocks without aggregate uncertainty as in Bewley (1977).

Preferences

Households are endowed with a fixed capacity of time $\bar{L}$ and derive utility from both consumption and leisure according to the period utility function:

$$u(c,\ell)=\frac{c^{1-\sigma_1}}{1-\sigma_1}-\psi\frac{\ell^{1-\sigma_2}}{1-\sigma_2} \tag{1}$$

Households have heterogeneous discount factors, i.e. they discount utility at rate $\beta \in \{\beta_1, \beta_2, \beta_3\}$. These discount factors are fixed during an individual's lifetime but may change across generations as they are transmitted from parent to child. This is captured in the transition process Γ_β.

When an old household dies she derives utility from the bequests she leaves to her children, giving her a voluntary bequest motive. Specifically, for leaving a bequest of size a, the old household receives utility according to "warm glow" preferences:

$$\phi(b(a))=\phi_1\left(1+\frac{b(a)}{\phi_2}\right)^{1-\sigma_3} \tag{2}$$

where $b(a)=a-\tau_b\max\{0,a-ex_b\}$ are proportional after-tax bequests. Notice that, as outlined in Castañeda *et al.* (2003), a proportional bequest tax appears to be in contrast to the U.S. tax code, which allows for a tax-exempt amount ex_b, followed by a progressive marginal tax rate. However, the effective marginal tax rate has been estimated to be significantly lower due to various loopholes (e.g. Aaron and Munnell (1992)) so that a proportional tax rate seems to be a better description of reality. For this reason, we assume the bequest tax to be proportional.

Technology

Households have two employment opportunities: they can either work as an employee or become an entrepreneur and run their own firm. Accordingly, households are endowed with two kinds of abilities: entrepreneurial ability, θ, and working ability, y, that evolve according to the stochastic processes Γ_θ and Γ_y respectively. Both young and middle-aged households decide whether to become an entrepreneur or a worker based on their endowment with these two abilities.

Entrepreneurs produce according to a technology whose returns are increasing in entrepreneurial ability. Specifically, their production function is dependent on entrepreneurial ability θ, invested capital k and hired labor n:

$$f(\theta, k, n) = \theta(k^{\gamma}(1+n)^{1-\gamma})^{\nu} \tag{3}$$

We assume $\nu < 1$, which implies that returns are decreasing in the size of the physical inputs, as managerial input gets spread over more resources. This means that the rate of return is endogenously determined for entrepreneurs and is a function of both the size of the project and her entrepreneurial ability. Also note that entrepreneurial output is still positive with $n = 0$ as the entrepreneur inputs some labor of her own, which reflects the large number of small entrepreneurial enterprises.

Following the approach laid out by Quadrini (2000) we also model a non-entrepreneurial sector. Firms in this sector are assumed to be sufficiently large to face no financial constraints and employ non-entrepreneurial workers according to a Cobb-Douglas production technology:

$$F(K_c, L_c) = AK_c^{\alpha}L_c^{1-\alpha} \tag{4}$$

where A is a technology parameter and K_c and L_c are capital and labor inputs. In both sectors capital depreciates at rate δ.

Credit Market Constraints

Working households can borrow up to their exogenous borrowing limit, $-\chi$. In contrast to this, entrepreneurs can only borrow through their firm. Specifically, we assume an exogenous borrowing constraint $k \leq \kappa a$, which can be micro-founded through the imperfect enforceability of contracts (see e.g. Moll (2014)).[3] In simple terms it states that entrepreneurs have to put up collateral for borrowing, which limits their borrowing capability. A consequence of the borrowing constraint is that not all projects will be fully funded, which results in different firm sizes even when entrepreneurial ability is held constant. Wealthier entrepreneurs will have larger firms not only because they can invest more of their own assets but also because the borrowing constraint is relaxed due to a higher collateral. Notice that, since an agent needs to forgo her wage as a worker to become an entrepreneur, she will only do so if her firm generates a sufficiently high return. This leads to an effective minimum firm size as households with assets below a certain threshold will decide to work independently of their entrepreneurial abilities. Note also that the optimal firm size for entrepreneurs is so large that even wealthy entrepreneurs will continue saving.

[3]In an earlier version of this paper, we implemented a full-fledged endogenous borrowing constraint, which yielded essentially identical results. Without loss of generality, we can thus simplify the problem substantially by using the exogenous borrowing constraint.

Households

Young households choose whether to be an entrepreneur or a worker. If they become an entrepreneur, they have to choose their consumption c, investment k, the number of labor hours hired n and savings a'. We assume that, for a given θ, every entrepreneur works the same (exogenously given) number of hours $\bar{\ell}(\theta)$. Entrepreneurs face (nonlinear) taxes on profits and capital income. In order to invest k, entrepreneurs borrow $k - a$ at the risk-free interest rate r. In contrast, workers choose consumption c, savings a', and the number of hours worked ℓ while facing taxes on capital and labor income. When households age, they receive a share $1 - \omega$ of their bequests b. The value function of a young household is defined as follows

$$V^y(a, y, \theta, \beta) = \max\{V_e^y(a, y, \theta, \beta), V_w^y(a, y, \theta, \beta)\} \tag{5}$$

where we denote the value function of a young entrepreneur with V_e^y and that of a young worker with V_w^y. They can be written as

$$\begin{aligned} V_e^y(a, y, \theta, \beta) = \max_{c,k,n,a'} \; & u(c, \bar{\ell}(\theta)) + \beta[\lambda_y \mathbb{E} V^y(a', y', \theta', \beta) + \\ & + (1 - \lambda_y)\mathbb{E} V^m(a' + (1 - \omega)b, y', \theta', \beta)] \end{aligned} \tag{6}$$

$$\text{s.t.} \begin{cases} a' = \Pi - \tau_y(\Pi) + a - (1 + \tau_c)c \\ k, a' \geq 0, \quad k \leq \kappa a \end{cases}$$

$$\begin{aligned} V_w^y(a, y, \theta, \beta) = \max_{c,a',\ell} \; & u(c, \ell) + \beta[\lambda_y \mathbb{E} V^y(a', y', \theta', \beta) + \\ & + (1 - \lambda_y)\mathbb{E} V^m(a' + (1 - \omega)b, y', \theta', \beta)] \end{aligned} \tag{7}$$

$$\text{s.t.} \begin{cases} a' = (1 + r)a - \tau_a(ra) + w\ell y - \tau_y(w\ell y) - (1 + \tau_c)c \\ a' \geq -\chi \end{cases}$$

where pre-tax profits are defined as $\Pi = \theta(k^\gamma(1 + n)^{1-\gamma})^v - \delta k - r(k - a) - wn$ and the expectations are taken with respect to (y', θ') conditional on (y, θ, β).

Middle-aged households can also be either entrepreneurs or workers and face a very similar problem to young households. The only differences is that they do not receive further bequests in their lives and are at a different stage in their life. Their value function is thus defined as follows:

$$V^w(a,y,\theta,\beta) = \max\{V_e^w(a,y,\theta,\beta), V_w^w(a,y,\theta,\beta)\} \tag{8}$$

where we denote the value function of a middle-aged entrepreneur with V_e^w and that of a middle-aged worker with V_w^m, which can again be written as

$$\begin{aligned} V_e^m(a,y,\theta,\beta) = \max_{c,k,n,a'} \; & u(c,\bar{\ell}(\theta)) + \beta[\lambda_m \mathbb{E}V^m(a',y',\theta',\beta) + \\ & + (1-\lambda_m)\mathbb{E}W(a',\theta',\beta)] \end{aligned} \tag{9}$$

$$\text{s.t.} \begin{cases} a' = \Pi - \tau_y(\Pi) + r - (1+\tau_c)c \\ k, a' \geq 0, \quad k \leq \kappa a \end{cases}$$

$$\begin{aligned} V_w^m(a,y,\theta,\beta) = \max_{c,a',\ell} \; & u(c,\ell) + \beta[\lambda_m \mathbb{E}V^m(a',y',\theta',\beta) + \\ & + (1-\lambda_m)\mathbb{E}W_r(a',\beta)] \end{aligned} \tag{10}$$

$$\text{s.t.} \begin{cases} a' = (1+r)a - \tau_a(ra) + w\ell y - \tau_y(w\ell y) - (1+\tau_c)c \\ a' \geq -\chi \end{cases}$$

Old households can either be retired or work as entrepreneurs. Retirement is permanent, that is retired households cannot start a new entrepreneurial activity. An entrepreneurial household on the other hand can keep her firm in old age. The rationale behind this modeling choice is that whilst it is unlikely for a retiree to start a firm, the data show a large number of entrepreneurs working into old age. We then have the value function of the old entrepreneur W:

$$W(a,\theta,\beta) = \max\{W_e(a,\theta,\beta), W_r(a,\beta)\} \tag{11}$$

where W_e and W_r are the values of staying an entrepreneur and retiring respectively. Note that y disappears as a state variable for the old as they cannot work outside of the entrepreneurial sector. Entrepreneurs can bequeath their entrepreneurial ability to their offspring according to the stochastic process Γ_θ^D. In contrast, the offspring of retired households get a draw from the unconditional distribution of θ. Finally, retired households receive a pension p. We then have

$$W_e(a,\theta,\beta) = \max_{c,k,n,a'} u(c,\bar{\ell}(\theta)) + \beta[\lambda_o \mathbb{E}W(a',\theta',\beta) + (1-\lambda_o)\phi(b(a'))] \tag{12}$$

$$\text{s.t} \begin{cases} a' = \Pi - \tau_y(\Pi) + a - (1+\tau_c)c \\ k, a' \geq 0, \quad k \leq \kappa a \end{cases}$$

$$W_r(a,\beta) = \max_{c,a'} u(c,0) + \beta[\lambda_o \mathbb{E}W_r(a',\beta) + (1-\lambda_o)\phi(b(a'))] \tag{13}$$

$$\text{s.t.} \begin{cases} a' = (1+r)a - \tau_a(ra) + p - (1+tau_c)c \\ a' \geq -\chi \end{cases}$$

where the expectations are taken with respect to θ' conditional on (θ,β).

3.b *Equilibrium*

We define a state vector $x = (t, s, a, y, \theta, \beta)$ where $t \in \{y, m, o\}$ indicates the age of the household and $s \in \{w, e\}$ her employment status, e being an entrepreneur and w a non-entrepreneurial household (i.e. worker or retiree). One can then define a transition function (see appendix A.1) that generates the probability distribution of x' conditional on x. A stationary equilibrium is given by

- An interest rate r and a wage w
- Government policies $(p, G, \tau_a(x), \tau_y(x), \tau_c(x), \tau_b, ex_b)$
- Policy functions $(c(x), k(x), \ell(x), n(c), a(x), D(x))$
- A distribution $m^*(x)$

such that:

- (r, w) satisfy firm FOCs:

$$r = F_K(K_c^*, L_c^*) - \delta \qquad w = F_L(K_c^*, L_c^*) \tag{14}$$

- given (r, w) and $(p, G, \tau_a(x), \tau_y(x), \tau_c(x), \tau_b, ex_b)$,
  - $(c(x), k(x), \ell(x), n(x), a(x), D(x))$ solve the household problem outlined above

– capital and labor markets clear:

$$K_c = \int_X a(x)\, dm^*(x) - \int_X k(x) dm^*(x) \tag{15}$$

$$L_c = \int_X y\ell(x|t \in \{y,m\}, s = w)\, dm^*(x) - \int_X n(x) dm^*(x) \tag{16}$$

- The government budget period clears every period:

$$\begin{aligned} G + \int_X p\mathbb{1}\{s = w, t = o\}\, dm^*(x) = & \int_X \tau_a(ra)\mathbb{1}\{s = w\}\, dm^*(x) + \\ & + \int_X \tau_c c(x)\, dm^*(x) + \\ & + \int_X \tau_y(w\ell(x)y)\mathbb{1}\{s = w\}\, dm^*(x) + \\ & + \int_X \tau_y(\Pi)\mathbb{1}\{s = e\}\, dm^*(x) + \\ & + \int_X \tau_b \lambda_o \max\{a(x) - ex_b, 0\}\mathbb{1}\{t = o\}\, dm^*(x) \end{aligned} \tag{17}$$

- The distribution m^* is the invariant distribution.

Here $D(x) \in \{e, w\}$ is the policy function for occupational choice and $X = \{y, m, o\} \times \{e, w\} \times A \times Y \times \Theta \times B$.

3.c *Modeling Choices and Extensions*

In the following we outline our reasoning behind the most important modeling steps and compare our choices to alternative specifications.

In order to fully represent the earnings risk channel introduced by Castañeda *et al.* (2003), we include an endogenous labor supply decision for working households. People in the awesome state will react to their earnings windfall by working a lot more due to both the high earnings and the low persistence of this state. Hence, in addition to the positive savings effects of higher y, there is an amplification mechanism due to the higher hours worked which further increases savings and inequality. Also, we have chosen not to give entrepreneurs a specific labor choice but rather a fixed workload, which depends on their entrepreneurial ability. The rationale is that if entrepreneurs had a labor choice, they would be able to dampen the adverse effect of shocks by working more and keeping their savings (and capital) high. For the same reason, despite including hired labor as an entrepreneurial production factor, we do not allow the entrepreneurs themselves to adjust their labor input.

In contrast to other models we do not include intergenerational transmission of working ability y. We have made this simplifying choice as this channel is not strictly necessary to answer our research question, given that the other models are already able to reproduce the wealth distribution.

Note that we do include transmission of entrepreneurial ability from old entrepreneurs to young. This is because the transmission of θ from parent to offspring represents the bequest of the firm itself. This bolsters firm size as wealthy offspring, who have the funds to run a big firm, also have a higher chance of being able to do so. Finally, since retired households do not have a firm, their earnings ability is not transmitted to their children who therefore receive a θ drawn from its unconditional distribution.

To include a voluntary bequest motive we have opted for introducing the warm glow bequest motive as in De Nardi (2004) instead of the full altruism model of Cagetti and De Nardi (2006). The reason behind this choice is that warm glow bequest motives are stronger and have the specific property of being able to model bequests as a luxury good.

We have included exogenous borrowing constraints for both households and entrepreneurs. As alluded to above, this simplifies the problem substantially compared to an endogenous borrowing constraint which we had in earlier versions of this paper, while changing none of the quantitative predictions. Furthermore, note that entrepreneurs can only borrow through the entrepreneurial borrowing channel and no longer through the one available to working households. We restrict the problem in this way do avoid the unnecessary complexity that would follow from allowing them to borrow through two different channels. Since it is not the point of this paper to explore the impact of different sources of financing on entrepreneurial activity, this is not a costly choice to make.

Concerning the earnings risk channel, we have opted for a linear earnings process with an awesome state as in Castañeda *et al.* (2003). This has the advantage of being both computationally feasible and relatively straightforward to calibrate. Recently, however, a range of new papers (Arellano 2014; Arellano, Blundell *et al.* 2017; Arellano and Bonhomme 2017a,b) has explored how to estimate non-linear earnings processes and have applied these estimation methods to models of consumption and wealth inequality (De Nardi, Fella *et al.* 2016).

4 Calibration

Our general model outlined above nests our four models of interest. This allows us to calibrate variables that are shared by the models to the same values, ensuring consistency between models. To calibrate the shared variables we use two approaches. First, for parameters that have well established values in the literature, such as the Gouveia and

Strauss (1994) tax schedule, we simply adopt those and provide the source. Second, for parameters that need to be calibrated to match specific moments, we do so in a basic form of our model with all the channels switched off. For example, we calibrate the disutility from work to match a work day of 8 hours or 1/3 of time available to the agents.

This leaves the model specific parameters. We calibrate these individually in each model with only the relevant channel being switched on. We then try to match the wealth distribution observed in the data as closely as possible with the available parameters of that channel. This is akin to viewing each channel as the ground truth and allowing it to explain the entire observed variation. In choosing this calibration strategy we level the playing field for the comparison of the competing channels.

A full overview of our calibration is provided in table 1. Let us first turn to the common parameters. We calibrate the probabilities of staying young, middle-aged and old, λ_y, λ_m and λ_o respectively, such that they match the typical life-cycle pattern. Specifically, a household remains young for an average of 15 years, middle-aged for 25 years and old for 15 years. We choose σ_1, the parameter of relative risk aversion, to be equal to 1.5, as estimated by Attanasio *et al.* (1999) and Gourinchas and Parker (2002). Similarly, we calibrate the analogous labor parameter σ_2 so that the inverse equals the Frisch elasticity of 1.5 as in Chang and Kim (2007). The preference weight for leisure ψ is calibrated such that the average hours worked are 1/3 as suggested by Cooley and Prescott (1995). In addition, we normalize the maximum amount of labor $\bar{L}$ to 1. We follow the standard business cycle literature by calibrating β such that the capital-output ratio is equal to three in line with Kaldor (1957). Of course, this calibration will be different in the heterogeneous discount factor model.

When calibrating the income process, we follow convention (e.g. Cagetti and De Nardi (2006)) in assuming that log-income follows an AR(1) process. As estimated by Storesletten *et al.* (2004) we assume the persistence parameter to be equal to 0.95 and choose the variance such that the Gini coefficient of income is equal to 0.38 as implied by the PSID. We then discretize this continuous process into a five-state Markov chain following Tauchen and Hussey (1991). Again, this calibration will differ in the earnings risk submodel. Calibration on the production side of the economy is standard and in line with most of the business cycle literature. We normalize the technology parameter A to 1 and set the capital share α in the non-entrepreneurial sector to 0.33 (see e.g. Gollin (2002)). Finally, we follow standard procedure in setting the depreciation rate δ to 0.06 as in Stokey and Rebelo (1995).

Next, we determine initial values for our policy parameters, which we are then going to vary in the subsequent policy experiments. We set the exogenous borrowing limit of workers to 0, thereby effectively banning borrowing of households who are not enterepreneurs. As in Kotlikoff, Smetters *et al.* (1999) we set the level of pensions to 40 percent of average yearly earnings. Concerning income tax rates, we follow Gouveia and Strauss (1994), who model the progressive income tax schedule. The exemption level from the bequest tax ex_b is calibrated such that it equals 40 times average yearly earnings, while we set the bequest tax rate τ_b to a value of 0.13 in line with the calibrations of Cagetti and De Nardi (2009). We follow De Nardi (2004) in setting the capital tax rate τ_a to 20% while letting the consumption tax rate τ_c be 11% as in Altig *et al.* (2001). Next, we calibrate the submodel-specific parameters. As discussed above, all of those parameters are calibrated on the wealth distribution. Specifically, we target the Gini coefficient as well as the wealth share of the top 1%, 5%, 20% and 40%. The results of this calibration can be seen in table 1.

4.a *Calibration of Channels*

As discuss above we calibrate the parameters relating to a specific channel on the wealth distribution, treating this channel effectively as the ground truth. Table 2 shows to what extend each channel is able to match the observed wealth distribution. The table summarizes the wealth distribution based on the wealth shares at different percentiles of the wealth distribution and the Gini coefficient. The first row of the table shows the data, as provided in Kuhn and Rios-Rull (2016). The remaining rows of the table compare the wealth distributions generated by the various model specifications. As a baseline we provide in the second row the "Standard" model, which has all of the channels switched off and was used to calibrate the common parameters, as outlined above. In this baseline configuration we already observe some inequality, which is due to the presence of earnings risk. However, the model's wealth distribution is far from the observed inequality in the data. Next we turn to "Earnings Risk" which includes the awesome state and therefore significantly more earnings variance. This channel is able to match all features of the observed wealth distribution closely. The same holds broadly for the "Entrepreneurship" model, which only has some minor deviations such as a wealth share of the top 5% that is too high. The "Bequest" channel on the other hand falls short significantly. It is neither able to match the observed Gini nor does it have high enough wealth concentration in the top wealth percentiles. The bequest motive, even in its strongest specification, i.e. warm glow bequest, is not able to generate enough right-tale inequality. Finally, heterogenous discount factors are once again able to to match the observed inequality, albeit overstating wealth shares for the top 5% and 10%.

Additionally, in Figure 1 we plot the average wealth in the three life-cycle phases for each model and compare it to the data. To simplify the comparison across models that have different overall levels of wealth and the data we normalize wealth to be 1 both at age 30 (in the data) and when young (in the model). As we can see the models all match

Table 1. Calibration

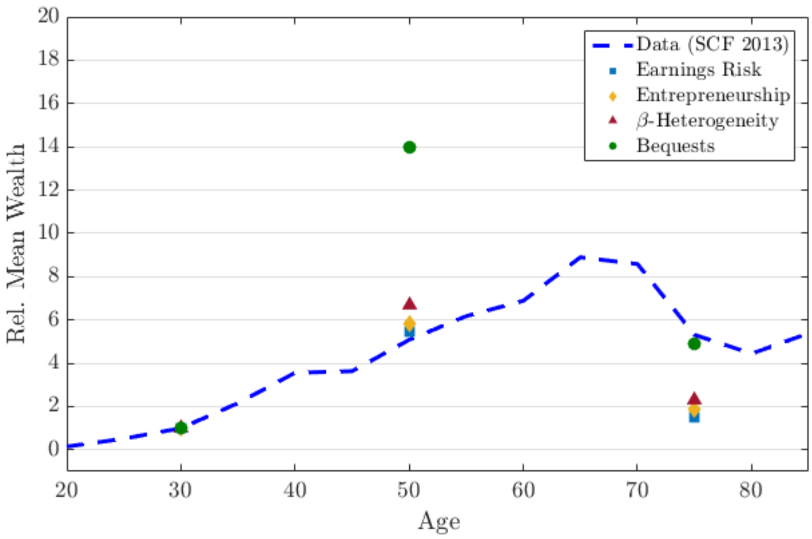
Parameter	Value	Source/Target
	Panel A: Common Parameters	
Household Parameters		
$(\lambda_y, \lambda_m, \lambda_o)$	$0.933, 0.96, 0.933$	Length of life-cycle phases
σ_1	1.5	Attanasio *et al.* (1999)
σ_2	1.5	1/Frisch elasticity Chang and Kim (2007)
ψ	10.5	Average hours = 1/3
$\bar{L}$	1	Normalization
β	0.94	K/Y=3
y	Appendix	see text
Γ_y	Appendix	see text
Firm Parameters		
A	1	Normalization
α	0.33	Gollin (2002)
δ	0.06	Stokey and Rebelo (1995)
Policy Parameters		
ex_b	endogenous	40 times average yearly earnings
τ_b	0.13	Cagetti and De Nardi (2009)
τ_a	0.2	De Nardi (2004)
τ_y		Gouveia and Strauss (1994)
τ_c	0.11	Altig *et al.* (2001)
τ_Π	0.22	U.S. Treasury
p	endogenous	40% replacement rate
	Panel B: Model-implied Parameters	
Earnings Risk		
y	Appendix	Model
Γ_y	Appendix	”
Entrepeneurship		
θ	[0, 0.99]	Wealth distribution
Γ_θ	Appendix	Fraction of entrepreneurs
$\bar{l}(\theta)$	[0, 0.2]	Wealth distribution
γ	0.84	Cagetti and De Nardi (2009)
κ	1.5	Cagetti and De Nardi (2009)
ν	0.88	Cagetti and De Nardi (2009)
$(\sigma^2_{e,y}, \sigma^2_{e,m})$	(0.061, 0.071)	Variance of earnings distribution
Bequests		
ϕ_1	-214	Wealth distribution
ϕ_2	24	”
Discount Factor Heterogeneity		
β	[0.8064, 0.9, 0.9936]]	”
Γ_β	Appendix	”

Table 2. Wealth Distribution

Source	Gini	Wealth in the Top				
		1%	5%	10%	20%	40%
Data	0.85	0.35	0.62	0.75	0.87	0.97
Standard	0.67	0.08	0.29	0.46	0.68	0.92
Earnings Risk	0.85	0.36	0.60	0.77	0.89	0.97
Entrepreneurship	0.86	0.34	0.66	0.80	0.91	0.98
Bequest	0.73	0.17	0.42	0.60	0.83	0.97
β-Heterogeneity	0.88	0.34	0.71	0.84	0.96	0.99

Source – Data: Kuhn and Rios-Rull (2016)

Figure 1. Age Profiles of Wealth



the stylized fact that average wealth goes up in middle age and that households then decumulate wealth late in life, although not quite as much as when they were young.

5 Results

Given the different models were all calibrated to match wealth inequality, all of the above results on the *level* of wealth inequality should come as no surprise. The more interesting question, which we address here, is whether it matters which of the competing models of inequality is adopted. Our strategy to do so is to analyze the steady state behavior of the individual models after a policy shock. In particular we hit each of the four considered models with the same policy shock and then compare the long-run responses of the economy to such policy change; in each policy experiment we look at the changes in top wealth inequality (top 1% and top 5%), overall wealth dispersion (Gini), and the overall

TABLE 3. Capital income tax, changes to baseline ($\tau_a = 0.2$)

	GDP	Top 1%	Top 5%	Gini
Panel A: $\tau_a = 0.3$				
β-Heterogeneity	-1.8%	-3.3%	-1.8%	-0.2%
Earnings Risk	-1.2%	1.7%	1.4%	0.6%
Entrepreneurship	0%	0%	0%	0.2%
Bequest	-1.3%	0.5%	1%	0.4%
Panel B: $\tau_a = 0.1$				
β-Heterogeneity	2.2%	6.5%	2.1%	0.4%
Earnings Risk	1.2%	-1.9%	-1.4 %	-0.7%
Entrepreneurship	0%	0%	0%	0%
Bequest	1.1%	0.3%	-1.1%	-0.3%

size of the economy. As the predictions of the competing channels vary significantly over the different policy shocks considered, in both magnitude and sign, we hereby show that the modeling environment does in fact matter for our key variable of interest, wealth inequality.

By showing that the responses to different policy shocks are consistently different across models we also show that, whenever a researcher chooses a specific mechanism to generate wealth inequality he is also implicitly taking a stance on what the response of the model will be to different policy shocks. Effectively, the way one decides to model wealth inequality matters for policy analysis.

In order to sharply contrast the models we deliberately pick large shocks. To this end we consider four policy experiments relating to two government variables of interest in this section, while appendix B includes the results for other policy experiments we did not find to be especially useful in distinguishing between models. First, we consider the effect of an increase in capital income taxes from 20% to 30%; to give a comprehensive overview we then also consider the counterpart of this policy experiment, a decrease of capital income taxes from 20% to 10%. Second, we consider the effects of changing the replacement rate of government-provided pensions. Differently from capital income taxes, changes in the replacement rate from 40% to 50% and 30% do not allow us to tease apart the different models.

5.a *Changes in capital income taxes*

We start by analyzing the effects of an increase in capital income taxes τ_a from 20% to 30% in Table 3. The effects on the models with preference heterogeneity are exactly as one would expect them to be: an increase in capital income taxes lowers the effective rate of return on wealth, which decreases households incentives to save throughout the distribution and reduces the overall size of the economy. On top of that, because saving-incentives are comparatively stronger for households at the top of the wealth distribution (patient, low-β households) these households dissave relatively more and wealth inequality is reduced. To put it differently, because households at the bottom of the wealth distribution behave almost like hand-to-mouth households independently of what their rate of return is, they don't really care about an increase in capital income taxes and don't change their decisions too much; on the other hand this is not true for richer households and the difference in the changes in saving rates along the wealth distribution implies an overall reduction in wealth inequality.

The results in the entrepreneurship model are a bit more surprising at first, as the model shows virtually no change after the policy shock. The reason why this is the case is that both the overall size of the economy and (especially) wealth inequality are determined by the size of the entrepreneurship sector and by the saving behavior of the entrepreneurs. However, because entrepreneurs' main saving motive is to grow their firm and the optimal firm size is both large and independent of capital income taxes, their trajectory is unchanged by a increase in capital income taxes. The only little action is at the bottom of the distribution, where workers now face a lower effective rate of return and their total wealth slightly decreases, which contributes to the small increase in the Gini coefficient.

Although for different reasons, both the bequest and earnings-risk models show an increase in wealth inequality coupled with a decrease in the overall size of the economy. On the one hand in the bequest model wealth inequality goes up because, while the incentive to save of the rich is unrelated to their effective rate of return (they still want to leave bequests), households at the bottom do not attach much weight to bequests due to their luxury-good nature and end up saving less. On the other hand in the earnings-risk model a decrease in the effective rate of return implies a relatively stronger precautionary-savings motive at the top than at the bottom, which also contributes to increased wealth inequality.

Concluding, a change in capital income taxes has very different implications for wealth inequality across different models. This has two main consequences: First, a researcher who wants to analyze the effects of a change in capital income taxes and "needs" to have a realistic wealth distribution will be also assuming his results away just by choosing what mechanism generates wealth inequality. Second, capital income taxes (and to some extent wealth taxes) represent a perfect identifying experiment to assess which of the channel "dominates" in the data. Namely, if after a reduction in capital income taxes we were to observe an increase in wealth inequality, we could only conclude the positive effect

Table 4. Pension reform, changes to baseline ($p = 0.4$)

	GDP	Top 1%	Top 5%	Gini
Panel A: $p = 0.5$				
β-Heterogeneity	-0.1%	14.9%	4.4%	2.4%
Earnings Risk	-1.6%	6.2%	5.3%	3.1%
Entrepreneurship	-3.1%	2.6%	1.8%	0.9%
Bequest	-1.5%	3%	3.5%	2.1%
Panel B: $p = 0.3$				
β-Heterogeneity	0.8%	-10.4%	-3.8%	-2.6%
Earnings Risk	2.0%	-7.1%	-6.3%	-3.8%
Entrepreneurship	3.5%	-2.7%	-1.8%	-0.9%
Bequest	2.0%	-3.4%	-4.2%	-2.6%

of the β-heterogeneity model is stronger than the effect of all other models combined. In fact, as we will see in section 5.c this is exactly the results in Jakobsen *et al.* (2020) who, after a reduction in wealth taxes, find a significant increase in wealth accumulation and wealth inequality.

5.b *Changes in pensions' replacement rate*

We hereby analyze the effects on wealth accumulation and inequality that follow a decrease (increase) in the replacement rate of government provided pensions from 40% to 30% (50%). As we can see in Table 4, unlike with capital income taxes, a pension reform has similar effects in all the different models (at least in sign): when the government reduces pensions households save more throughout the distribution and the size of the economy increases. However, this is especially true for households at the bottom of the wealth distribution who need to rely more on their own savings to go through the last phase of their life. Consequently inequality decreases strongly in all considered models. We can therefore conclude that the policymaker who wishes to analyze the effects of a pension reform in the context of a model with a realistic wealth distribution can freely choose which channel is the most appropriate as that will not imply significant differences in the response of the wealth distribution to the policy shock.

5.c *Channel identification*

In the previous two subsections we analyzed the long-run responses of each model to different policy experiments and found that, while some policy changes induce very different changes depending on which model we are looking at, others induce exactly the same (qualitative) response of wealth inequality. Would it not be be great if we could find, in the data, a natural policy experiment that we could replicate in our models and that would also give us different response in each model? Fortunately it turns out that such a natural experiment exists and has already been thoroughly studied in Jakobsen *et al.* (2020). Even better, that experiment uniquely pins down which of the four channels analyzed here is the most relevant one.

Using administrative data from Denmark, Jakobsen *et al.* (2020) study the effects of a change in wealth taxation on wealth accumulation and find that a reduction in wealth taxes implies a significant increase in wealth accumulation. Given the very well-known result that capital income taxes and wealth taxes are equivalent when the rate of return is the same across households Guvenen, Kambourov *et al.* (2019), it immediately follows that in our model a reduction in wealth taxes is the same as a reduction in capital income taxes, exactly the experiment we ran in section 5.a. On top of that the only model capable of generating a significant increase in wealth inequality after a reduction in capital income taxes was the β-heterogeneity model. We conclude that, among all channels considered, heterogeneity in discount factors is most likely the dominant channel. [4]

6 Conclusion

A long literature has documented a number of channels that is capable of generating wealth inequality at the top. In particular, earnings risk, a voluntary bequest motive, entrepreneurship and discount factor heterogeneity are ways to account for this phenomenon. This paper aims to contribute by teasing these channels apart. Specifically, we do so by building a general model that incorporates all those channels to quantitatively examine how the implications of each model differ in response to a set of policy experiments. We find significant differences in the model predictions in response to key policy experiments.

Among the policy experiments we run, the most important one is a change in capital gains taxation. This experiment provides starkly different predictions for the heterogeneous returns channel compared to all others. In particular, it predicts that wealth inequality increases in response to a decrease in capital gains taxes and vice versa. In contrast, for all other channels, the model predicts either a decrease or no response at all. This sharp contrast in predictions allows us to draw on the empirical results in Jakobsen et al. (2020) to identify the dominant channel of wealth inequality.

[4] This is also in line with the recent evidence on heterogeneity in rates of return provided by Bach *et al.* (2020), Fagereng *et al.* (2020), and M. Smith *et al.* (2019). In fact, if one looks at the standard Euler equation for all these models the rate of return on wealth r and the discount factor β always enter multiplicatively and, from a modeling perspective, whether save more because they have a higher discount factor or a higher interest rate is observationally equivalent. This means one can also interpret the β-heterogeneity model as both a reduced form way of modeling heterogeneity in rates of return.

In particular, those authors find that a decrease in wealth taxation indeed leads to an increase in wealth inequality, as predicted solely by the heterogeneous returns channel. We therefore conclude that heterogeneous returns, as modeled by heterogeneous discount factors, is the dominant channel in generating wealth inequality.

This result aids both academics and policy makers in their study and understanding of, as well as their response to wealth inequality. In particular, policy makers concerned with increasing wealth inequality and the inherent tradeoffs of addressing it, can focus on measures targeted at rate of return heterogeneity. For instance, based on this result it appears that high costs of investment for unsophisticated investors might be a driving force in increased wealth inequality. For academics, this result has two implications. First, this evidence for the importance of heterogeneous returns invites further research into the origins and mechanisms involved. Second, researchers studying questions related to wealth inequality can use heterogeneous discount factors as a simple modeling device to capture the first-order effects on the wealth distribution.

Appendix A

A.1 *Invariant distribution*

The transition function is the probability distribution of $x' \in \tilde{X}$ conditional on $x \in X$, i.e. it maps from the set of this period's state variables X into the set of next period's state variables $\tilde{X}$. Specifically, X and $\tilde{X}$ are defined as follows:

$$
\begin{aligned}
X &= \{y, m, o\} \times \{w, e\} \times A \times Y \times \Theta \times B \\
\tilde{X} &= \mathscr{P}(\{y, m, o\}) \times \mathscr{P}(\{w, e\}) \times \mathscr{B}(A) \times \mathscr{B}(Y) \times \mathscr{B}(\Theta) \times \mathscr{B}(B)
\end{aligned}
$$

where A is the asset space, Y the income space, Θ the space of entrepreneurial ability, B the space of discount factors, $\mathscr{B}(A)$ the Borel σ-algebra of A and $\mathscr{P}$ is the cardinal set.

The following operator maps distributions from X to $\tilde{X}$:

$$
(R_M m)(x') = \int_X M(x, x')\, dm(x) \qquad \forall\, x' \in \tilde{X}
$$

where the transition function $M(x, x')$ is defined as follows:

$$
\begin{aligned}
M(x, x') = \mathbb{1}\{a' = a(x)\}\Bigg\{ & \mathbb{1}\{t = \text{young}\}\lambda_y \Gamma_y(y'|y)\Gamma_\theta(\theta'|\theta)\mathbb{1}\{s' = D(a', y', \theta', \beta, \text{young})\} + \\
& + \Big(\mathbb{1}\{t = \text{young}\}(1 - \lambda_y) + \mathbb{1}\{t = \text{middle}\}\lambda_m\Big)\Gamma_y(y'|y)\Gamma_\theta(\theta'|\theta)\mathbb{1}\{s' = D(a', y', \theta', \beta, \text{middle})\} \\
& + \Big(\mathbb{1}\{t = \text{middle}\}(1 - \lambda_m) + \mathbb{1}\{t = \text{old}\}\lambda_o\Big) \times \\
& \times \Big[\mathbb{1}\{s = e\}\Big(\Gamma_\theta(\theta'|\theta)\mathbb{1}\{D(a', y', \theta', \beta, \text{old}) = e\} + \mathbb{1}\{D(a', y', \theta', \beta, \text{old}) = w\}\Big) + \mathbb{1}\{s = w\}\Big] + \\
& + \mathbb{1}\{t = \text{old}\}(1 - \lambda_o)\Gamma_y(y')\Gamma_\beta(\beta'|\beta)\mathbb{1}\{s' = D(a', y', \theta', \beta', \text{young})\} \times \\
& \times \Big(\mathbb{1}\{s = e\}\Gamma_\theta^D(\theta'|\theta) + \mathbb{1}\{s = w\}\Gamma_\theta(\theta')\Big)\Bigg\}
\end{aligned}
$$

The invariant distribution m^* is then the fixed point to:

$$
m^* = R_M\, m^*
$$

A.2 *OLG model*

Environment

In this section we outline an alternative model that we discussed. It is an OLG model in the tradition of De Nardi (2004). People live for 14 periods, where each period represents five years. People enter the model in period one at age 20, and have a child in period two.

The child comes of age in period six with zero assets and observes y or θ that the parent has at that time. We combine this in a variable sp that takes the value of a working parent's y and an entrepreneurial parent's θ. This information is used to compute the bequests that a child expects to receive in case the parent dies. sp stays constant until the parent dies at which point it becomes equal to zero.

From period nine onwards agents face an age-dependent death probability s_t. In period ten working agents retire and die for sure by then end of period 14. The entrepreneurial choice formulation as well as the remaining environment are the same as in section 3. The value function for young households is defined as follows:

$$V(t,a,y,sp,\beta,\theta) = \max\{V_e(t,a,y,sp,\beta,\theta), V_w(t,a,y,sp,\beta,\theta)\} \tag{A.1}$$

where the value functions of entrepreneurs, workers and across the various ages are the following:

<u>$t = 1$ to $t = 3$</u>

$$V_e(t,a,y,sp,\beta,\theta) = \max_{c,a',k} u(c,\bar{l}(\theta)) + \beta\mathbb{E}_t V(t+1,a',y',sp,\beta',\theta') \tag{A.2}$$

$$\text{s.t.} \begin{cases} a' = \Pi - \tau_\Pi(\Pi) + (1+r)a - \tau_a(ra) - c - \tau_c(c) \\ c + \tau_c(c) \leq \Pi - \tau_\Pi(\Pi) + (1+r)a - \tau_a(ra) \\ V_w(t, f \cdot k, y, sp, \beta, \theta) \leq u(c,\bar{l}(\theta)) + \beta\mathbb{E}_t V(t+1,a',y',yp,\beta',\theta') \end{cases}$$

$$V_w(t,a,y,sp,\beta,\theta) = \max_{c,a',l} u(c,l) + \beta\mathbb{E}_t V(t+1,a',y',sp,\beta',\theta') \tag{A.3}$$

$$\text{s.t.} \begin{cases} a' = (1+r)a - \tau_a(ra) + wly - \tau_l(wly) - c - \tau_c(c) \\ c + \tau_c(c) \leq (1+r)a - \tau_a(ra) + wly - \tau_l(wly) + \chi \\ l \in [0, \bar{L}] \end{cases}$$

$\underline{t = 4 \text{ to } t = 8}$

$$V_e(t,a,y,sp,\beta,\theta) = \max_{c,a',k} u(c,\bar{l}(\theta)) + \beta \mathbb{E}_t V(t+1,a',y',sp',\beta',\theta') \tag{A.4}$$

$$\text{s.t.} \begin{cases} a' = & \Pi - \tau_\Pi(\Pi) + (1+r)a - \tau_a(ra) - c - \tau_c(c) + \\ & + b'\mathbb{1}\{sp > 0\}\mathbb{1}\{sp' = 0\} \\ c + \tau_c(c) \leq \Pi - \tau_\Pi(\Pi) + (1+r)a - \tau_a(ra) \\ V_w(t, f\cdot k, y, sp, \beta, \theta) \leq u(c,\bar{l}(\theta)) + \beta \mathbb{E}_t V(t+1,a',y',yp',\beta',\theta') \\ sp' = \begin{cases} sp & \text{with probability } s_{t+5} \\ 0 & \text{with probability } 1 - s_{t+5} \end{cases} \end{cases}$$

$$V_w(t,a,y,sp,\beta,\theta) = \max_{c,a',l} u(c,l) + \beta \mathbb{E}_t V(t+1,a',y',sp',\beta',\theta') \tag{A.5}$$

$$\text{s.t.} \begin{cases} a' = & (1+r)a - \tau_a(ra) + wly - \tau_l(wly) - c - \tau_c(c) + \\ & + b'\mathbb{1}\{sp > 0\}\mathbb{1}\{sp' = 0\} \\ c + \tau_c(c) \leq (1+r)a - \tau_a(ra) + wly - \tau_l(wly) + \chi \\ l \in [0, \bar{L}] \\ sp' = \begin{cases} sp & \text{with probability } s_{t+5} \\ 0 & \text{with probability } 1 - s_{t+5} \end{cases} \end{cases}$$

$\underline{t = 9}$

$$\begin{aligned} V_e(t,a,y,sp,\beta,\theta) = \max_{c,a',k} & u(c,\bar{l}(\theta)) + \beta s_t \mathbb{E}_t W(t+1,a',\beta',\theta') + \\ & + (1-s_t)\phi(b(a')) \end{aligned} \tag{A.6}$$

$$\text{s.t.} \begin{cases} a' = & \Pi - \tau_\Pi(\Pi) + (1+r)a - \tau_a(ra) - c - \tau_c(c) + \\ & + b'\mathbb{1}\{sp > 0\}\mathbb{1}\{sp' = 0\} \\ c + \tau_c(c) \leq \Pi - \tau_\Pi(\Pi) + (1+r)a - \tau_a(ra) \\ V_w(t, f\cdot k, y, sp, \beta, \theta) \leq & u(c,\bar{l}(\theta)) + s_t \beta \mathbb{E}_t W(t+1,a',\beta',\theta') + \\ & + (1-s_t)\phi(b(a')) \\ sp' = \begin{cases} sp & \text{with probability } s_{t+5} \\ 0 & \text{with probability } 1 - s_{t+5} \end{cases} \end{cases}$$

$$
\begin{aligned}
V_w(t,a,y,sp,\beta,\theta) = &\max_{c,a',l} u(c,l) + \beta s_t \mathbb{E}_t W_r(t+1,a',\beta') + \\
&+ (1-s_t)\phi(b(a')) \\
&\text{s.t.} \begin{cases}
a' = & (1+r)a - \tau_a(ra) + wly - \tau_l(wly) - c - \tau_c(c) + \\
& \quad + b'\mathbb{1}\{sp>0\}\mathbb{1}\{sp'=0\} \\
c + \tau_c(c) \leq (1+r)a - \tau_a(ra) + wly - \tau_l(wly) + \chi \\
l \in [0,\bar{L}] \\
sp' = \begin{cases} sp & \text{with probability } s_{t+5} \\ 0 & \text{with probability } 1 - s_{t+5} \end{cases}
\end{cases}
\end{aligned}
\tag{A.7}
$$

For old households we have:

$$
W(t,a,\beta,\theta) = \max\{W_e(t,a,\beta,\theta), W_r(t,a,\beta)\} \tag{A.8}
$$

where the entrepreneurial and retired value functions are given by:

<u>$t = 10$ to $t = 14$</u>

$$
\begin{aligned}
W_e(t,a,\beta,\theta) = &\max_{c,a',k} u(c,\bar{l}(\theta)) + \beta s_t \mathbb{E}_t W(t+1,a',\beta',\theta') + \\
&+ (1-s_t)\phi(b(a')) \\
&\text{s.t.} \begin{cases}
a' = \Pi - \tau_\Pi(\Pi) + (1+r)a - \tau_a(ra) - c - \tau_c(c) \\
c + \tau_c(c) \leq \Pi - \tau_\Pi(\Pi) + (1+r)a - \tau_a(ra) \\
W_r(t,a,\beta) \leq & u(c,\bar{l}(\theta)) + \beta s_t \mathbb{E}_t W(t+1,a',\beta',\theta') + \\
& \quad + (1-s_t)\phi(b(a')
\end{cases}
\end{aligned}
\tag{A.9}
$$

$$
\begin{aligned}
W_r(t,a,\beta) = &\max_{c,a',l} u(c,l) + \beta s_t \mathbb{E}_t W_r(t+1,a',\beta') + (1-s_t)\phi(b(a')) \\
&\text{s.t.} \begin{cases}
a' = (1+r)a + p - c - \tau_c(c) \\
c + \tau_c(c) \leq (1+r)a - \tau_a(ra) + p + \chi \\
l \in [0,\bar{L}]
\end{cases}
\end{aligned}
\tag{A.10}
$$

As a terminal condition we impose:

$$
W(15,a,\beta,\theta) = \phi(b(a)) \tag{A.11}
$$

In the above, the functions Π, b and ϕ are the same as in section 3.

Equilibrium

We define a state vector $x = (t, s, a, y, \theta, \beta)$ where $t \in \{\mathbb{N} \cap [1, 14]\}$ indicates the age of the household and $s \in \{w, e\}$ the employment status where e indicates an entrepreneurial and w a non-entrepreneurial (ie working or retired) household. One can then define a transition function that generates the probability distribution of x' conditional on x. A stationary equilibrium is given by

- An interest rate r and a wage w
- Government policies $(p, G, \tau_a(x), \tau_\Pi(x), \tau_y(x), \tau_c(x), \tau_b, ex_b)$
- Policy functions $(c(x), k(x), l(x), a(x), D(x))$
- A distribution $m^*(x)$

such that:

- (r, w) satisfy firm FOCs:

$$r = F_K(K_c^*, L_c^*) - \delta \qquad w = F_L(K_c^*, L_c^*) \tag{A.12}$$

- given (r, w) and $(p, G, \tau_a(x), \tau_\Pi(x), \tau_y(x), \tau_c(x), \tau_b, ex_b)$,
  - $(c(x), k(x), l(x), a(x), D(x))$ solve the household problem outlined above
  - capital and labor markets clear:

$$K = \int_X a(x)\, dm^*(x) - \int_X k(x) dm^*(x) \tag{A.13}$$

$$L = \int_X l(x | t \leq 9, s = w)\, dm^*(x) \tag{A.14}$$

- The government budget period clears every period:

$$
\begin{aligned}
G + \int_X p\mathbb{1}\{s = w, t \geq 10\}\, dm^*(x) = &\int_X \tau_a(ra)\, dm^*(x) + \\
&+ \int_X \tau_c(x)\, dm^*(x) + \\
&+ \int_X \tau_\Pi(\Pi(x))\mathbb{1}\{s = e\}\, dm^*(x) + \\
&+ \int_X \tau_y(wl(x)y)\mathbb{1}\{s = w, t \leq 9\}\, dm^*(x) + \\
&+ \int_X \tau_b \max\{a(x) - ex_b, 0\}\mathbb{1}\{t \geq 10\}\lambda_o\, dm^*(x)
\end{aligned} \tag{A.15}
$$

- The family of expected bequest distributions $\mu_b(x)$ is consistent with the bequests that are actually left by the parents.

- The distribution m^* is the invariant distribution.

As in section 3, $D(x) \in \{e, w\}$ is the policy function for occupational choice and $X = \{\text{young}, \text{old}\} \times \{e, w\} \times A \times Y \times \Theta \times B$.

Appendix B

B.1 *Algorithm*

To solve our infinite horizon model we employ a value function iteration embedded within two loops that are explained below. The first loop iterates on the guesses for the interest rate. To this end we use the policy functions combined with the transition probabilities for the exogenous states to construct the transition matrix. Note that this matrix is sparse and so we benefit from a decomposed allocation to memory, i.e. only save the coordinates as well as the values of the entries to a set of vectors. We can then iterate an initial guess for the invariant distribution to arrive at the distribution in equilibrium. This then allows us to aggregate. We can now calculate the implied interest rate from the investment in the non-entrepreneurial sector. The interest rate loop runs until we achieve convergence between the given guess and the implied interest rate, where the implied interest rate is given by the FOC of the non-entrepreneurial sector.

The second loop before the value function iteration is dedicated to finding the borrowing constraints for the entrepreneurs. Here we allow them to first borrow the maximal amount on the capital grid. Then after each complete value function iteration we compare the utility from defaulting to that of honoring the contract and adjust the borrowing constraint accordingly so that we do not have default in equilibrium. We iterate until the borrowing constraints is not violated in any state.

B.2 *Calibration Details*

This section provides the remaining calibrated transition matrices that were not included in the calibration section for space considerations: Let us start with the **standard calibration for the earnings process** y and its transition matrix Γ_y:

$$y = \begin{bmatrix} .2468 & .4473 & .7654 & 1.3097 & 2.3742 \end{bmatrix}$$
$$\Gamma_y = \begin{bmatrix} .7376 & .2473 & .0150 & .0002 & .0000 \\ .1947 & .5555 & .2328 & .0169 & .0001 \\ .0113 & .2221 & .5333 & .2221 & .0113 \\ .0001 & .0169 & .2328 & .5555 & .1947 \\ .0000 & .0002 & .0150 & .2473 & .7376 \end{bmatrix}$$

Next, we provide the calibration for the **earnings process in the earnings-risk submodel**:

$$y = \begin{bmatrix} 1.0000 & 3.1500 & 9.7800 & 1061.0000 \end{bmatrix}$$
$$\Gamma_y = \begin{bmatrix} 0.96234 & 0.01140 & 0.00390 & 0.00006 \\ 0.03070 & 0.94330 & 0.00370 & 0.00000 \\ 0.01500 & 0.00430 & 0.95820 & 0.00200 \\ 0.10660 & 0.00490 & 0.06110 & 0.80510 \end{bmatrix}$$

We now turn to the transition matrix for **entrepreneurial ability** Γ_θ, which in this calibration is set equal to the intergenerational transition matrix of entrepreneurial ability Γ_θ^D:

$$\Gamma_\theta = \Gamma_\theta^D = \begin{bmatrix} 0.964 & 0.036 \\ 0.206 & 0.794 \end{bmatrix}$$

Finally, the intergenerational transition matrix for β in the submodel with **discount factor heterogeneity** is the following:

$$\Gamma_\beta = \begin{bmatrix} 0.600 & 0.400 & 0.000 \\ 0.100 & 0.875 & 0.025 \\ 0.000 & 0.300 & 0.700 \end{bmatrix}$$

B.3 *Replications*

We have replicated De Nardi (2004) as this is the crucial paper in the bequest literature. It illustrates the strength of the voluntary bequest motive while also being a full multi-period OLG model. This sets it apart from the other papers we replicated and was very useful for determining the fundamental structure of the model in this paper.

The results of such replication can be found in the first four lines of **??**. The first and third lines depict the results for the U.S. and Swedish calibrations, respectively. We added the latter in order to demonstrate the robustness of the code. Note that the replication is exact for the amount of wealth held by the various sub-groups as well as for the fraction of people with zero or negative wealth.

We choose Cagetti and De Nardi (2006) as our replication paper for the entrepreneurial channel as it is able to successfully generate high wealth concentration in the top one percent through a large optimal firm size and borrowing constraints for the entrepreneurs. Panel A of **??** shows the results of our replication. Note that the code is sensitive to grid specification and grid size, leading to slight deviations from the results presented in the paper.

Finally, we replicated Castañeda *et al.* (2003), which successfully reproduces the distributions of earnings and wealth of U.S. data by means of an unorthodox calibration of the earnings process. The replication managed to closely recreate all of the main tables in the original paper. In particular, panel B of **??** includes the replication results of table 7 in Castañeda *et al.* (2003), which shows that all figures are virtually identical. All small differences can likely be attributed to two main differences: as the authors do not provide many details on the solution algorithm, this is likely to be very different; also, due to numerical problems, the model was solved with a proportional income tax scheme, rather than a progressive scheme as in the original paper.

B.4 *Other policy experiments*

In tables B.1 and B.2 we include the results from other policy experiments which we decided to leave out of the main text because the results don't look as promising as for the other policy experiments. We still plan to analyze these further to check whether we can some more insight.

B.5 *Future policy experiments*

In this section, we outline the expected qualitiative behavior of policy experiments that we are planning to conduct in the future.

Table B.1. Capital Income Taxes

Model	Gini	Wealth in the Top				Entrep.	B/A	K/Y	Average Hours
		1%	5%	20%	40%				
					Panel A: $\tau_a = 0$				
Voluntary Bequests									
Results	0.75	0.45	0.56	0.78	0.93	-	0.01	4.22	0.36
%Δ to baseline	7.80	32.10	19.20	6.70	2.10	-	4.30	10.40	0.60
Entrepreneurship									
Results	0.83	0.3	0.63	0.89	0.97	0.1	0.01	5.21	0.31
%Δ to baseline	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Beta Heterogeneity									
Results	0.84	0.35	0.65	0.89	0.98	-	0.01	3.46	0.38
%Δ to baseline	1.30	9.90	4.10	0.70	0.10	-	2.10	9.70	-1.40
Awesome State									
Results	0.81	0.36	0.49	0.83	0.97	-	0.01	3.66	0.31
%Δ to baseline	-1.10	-3.00	-2.20	-1.60	-0.50	-	2.00	8.00	-2.40
					Panel B: $\tau_a = 0.4$				
Voluntary Bequests									
Results	0.72	0.37	0.5	0.75	0.92	-	0.01	3.53	0.37
%Δ to baseline	2.70	8.40	5.40	2.60	1.10	-	-3.80	-7.60	3.30
Entrepreneurship									
Results	0.83	0.3	0.63	0.89	0.97	0.1	0.01	5.21	0.31
%Δ to baseline	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Beta Heterogeneity									
Results	0.81	0.28	0.59	0.88	0.98	-	0.01	2.78	0.39
%Δ to baseline	-1.50	-11.50	-5.00	-1.00	-0.10	-	-2.90	-11.90	2.20
Awesome State									
Results	0.82	0.38	0.52	0.85	0.98	-	0.01	3.08	0.33
%Δ to baseline	1.20	3.70	2.60	1.50	0.50	-	-2.30	-9.00	2.90

Table B.2. Consumption Taxes

Model	Gini	Wealth in the Top				Entrep.	B/A	K/Y	Average Hours
		1%	5%	20%	40%				
					Panel A: $\tau_c = 0$				
Voluntary Bequests									
Results	0.74	0.43	0.54	0.77	0.93	-	0.01	3.93	0.36
%Δ to baseline	6.30	24.10	14.70	5.50	1.80	-	1.50	2.70	-0.50
Entrepreneurship									
Results	0.83	0.3	0.63	0.89	0.97	0.1	0.01	5.21	0.31
%Δ to baseline	0.00	-0.10	0.00	-0.10	0.00	-1.60	-0.90	0.00	-2.20
Beta Heterogeneity									
Results	0.83	0.32	0.63	0.88	0.98	-	0.01	3.15	0.38
%Δ to baseline	0.00	0.10	0.10	0.00	0.00	-	0.10	0.00	-2.40
Awesome State									
Results	0.81	0.37	0.5	0.84	0.98	-	0.01	3.39	0.31
%Δ to baseline	0.00	-0.20	0.10	0.10	0.00	-	0.00	0.00	-2.30
					Panel B: $\tau_c = 0.22$				
Voluntary Bequests									
Results	0.69	0.32	0.45	0.73	0.91	-	0.01	3.77	0.37
%Δ to baseline	-1.50	-6.00	-3.70	-1.30	-0.40	-	-1.60	-1.40	2.00
Entrepreneurship									
Results	0.83	0.3	0.63	0.89	0.97	0.1	0.01	5.2	0.32
%Δ to baseline	0.20	0.90	0.40	0.10	0.00	0.20	0.80	-0.10	2.30
Beta Heterogeneity									
Results	0.83	0.32	0.63	0.88	0.98	-	0.01	3.15	0.4
%Δ to baseline	0.00	0.10	0.10	0.00	0.00	-	0.00	0.00	2.20
Awesome State									
Results	0.81	0.37	0.5	0.84	0.98	-	0.01	3.39	0.33
%Δ to baseline	-0.10	-0.20	-0.10	-0.10	0.00	-	0.10	0.20	2.20

Tightening of the Borrowing constraint A shock to the borrowing constraints in the model directly affects the savings responses of the agents. The specific experiment in question is to simultaneously shock the parameters that govern agents' borrowing. One source of such a shock could for example be a change to chapter 13 of the bankruptcy code, or a possible change in the lending standards that regulators impose.

Let us first consider the entrepreneurship model. Here we have two groups of agents: workers and the entrepreneurs. For entrepreneurs the key borrowing parameter is the retained share of capital after default, f, while for households it is the exogenous borrowing limit, χ (or alternatively, in the presence of endogenous borrowing constraints as in Zhang (1997), the value of autarky). What mechanism drives the response of each group of agents?

Consider entrepreneurs first: when the borrowing limit tightens this changes the effective size of their company, conditional on them being at or close to their borrowing constraint. Hence they suffer a negative income shock. Note that as there are decreasing returns to capital a poor entrepreneur has a higher savings rate than his wealthier counterpart ceteris paribus. These factors combined lead to a high savings rate among entrepreneurs. Moreover, Cagetti and De Nardi (2006) have shown that inequality in this context decreases as f increases. The reason here is that entrepreneurial firm size decreases under more stringent borrowing constraints. Now let us turn to poor workers: As their borrowing constraint tightens they are forced to save more. Hence, we expect an unambiguous decrease in inequality and an increase of average savings rate in the model of entrepreneurship.

MIT Shock to productivity Let us now turn to an issue considered by Bassetto *et al.* (2015). This experiment aims to separate the entrepreneurship channels from the others. In the context of the restricted entrepreneurial model, this shock takes the shape of both a TFP shock to A, non-entrepreneurship productivity, and θ, entrepreneurial ability. What do we expect the effect of this drop in productivity to be? Entrepreneurial firms shrink; as this affects their borrowing limit, it takes them a long time to re-build their firms, causing the shock to be persistent. This implies a high savings rate for the entrepreneurs and, as the richest entrepreneurs stand to lose the most, this force leads to a persistent decline in inequality.

The entrepreneurship model shares the behavior of working households with the other restricted models. Their behavior is driven by multiple considerations. On the one hand, they have a negative income shock inducing an increase of savings. On the other, they are also in a transitory bad state which should induce higher spending. Consequently, the experiment at hand tries to tease the models apart by comparing the aggregate saving rates as well as those of the wealthy.

Tax Rebates One of the features of the preference heterogeneity channel is that it creates differential Marginal Propensities to Consume (MPC). We can exploit this fact to differentiate this channel. In particular, we can analyze the effect of tax rebates that come in the form of a lump sum shock to households. Since we expect a mass of low beta households to be close to the exogenous borrowing constraint, a shock to their income would generate a relatively high MPC among them. This could drive a high average MPC among the low wealth households in the discount factor heterogeneity model and would allow us to set it apart from the others. More generally we expect the profile of MPCs over the asset space to vary significantly, with high beta types on average being in the right tail of the distribution. Hence, we can also compare the entire profile of MPCs among the various restricted models. We predict that the MPC profile of the heterogeneous beta model looks significantly different to that of the other models.

Bibliography

Aaron, H. J. and A. H. Munnell (June 1992). “Reassessing the Role for Wealth Transfer Taxes”. In: *National Tax Journal* 45, pp. 119–43.

Aiyagari, S. R. (Aug. 1994). “Uninsured Idiosyncratic Risk and Aggregate Saving”. In: *Quarterly Journal of Economics* 109(3), pp. 119–43.

Altig, D., A. J. Auerbach, L. J. Kotlikoff, K. A. Smetters, and J. Walliser (2001). “Simulating Fundamental Tax Reform in the United States”. In: *American Economic Review* 91(3), pp. 574–595.

Arellano, M. (2014). “Uncertainty, Persistence, and Heterogeneity: A Panel Data Perspective”. In: *Journal of the European Economic Association* 12(5), pp. 1127–1153.

Arellano, M., R. Blundell, and S. Bonhomme (2017). “Earnings and Consumption Dynamics: A Nonlinear Panel Data Framework”. In: *Econometrica* 85(3), pp. 693–734.

Arellano, M. and S. Bonhomme (2017a). “Nonlinear Panel Data Methods for Dynamic Heterogeneous Agent Models”. In: *Annual Review of Economics* 9, pp. 471–496.

— (2017b). “Quantile Selection Models with an Application to Understanding Changes in Wage Inequality”. In: *Econometrica* 85(1), pp. 1–28.

Attanasio, O. P., J. Banks, C. Meghir, and G. Weber (Jan. 1999). “Humps and Bumps in Lifetime Consumption”. In: *Journal of Business and Economic Statistics* 17, pp. 22–35.

Bach, L., L. E. Calvet, and P. Sodini (Sept. 2020). “Rich Pickings? Risk, Return, and Skill in Household Wealth”. en. In: *American Economic Review* 110(9), pp. 2703–2747.

Bassetto, M., M. Cagetti, and M. De Nardi (2015). “Credit Crunches and Credit Allocation in a Model of Entrepreneurship”. In: *Review of Economic Dynamics* 18(1), pp. 53–76.

Becker, G. S. and N. Tomes (Dec. 1979). “An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility”. In: *Journal of Political Economy* 87(6), pp. 1153–89.

Benhabib, J. and A. Bisin (2017). “Skewed Wealth Distributions: Theory and Empirics”. In: *Journal of Economic Literature* Forthcoming.

Benhabib, J., A. Bisin, and M. Luo (2016). “Wealth Distribution and Social Mobility in the US: A Quantitative Approach”. In: *Mimeo.*

Bewley, T. F. (Dec. 1977). "The Permanent Income Hypothesis: A Theoretical Formulation". In: *Journal of Economic Theory* 16, pp. 252–92.

Blanchard, O. J. (Apr. 1985). "Debt, Deficits, and Finite Horizons". In: *Journal of Political Economy* 93, pp. 223–47.

Cagetti, M. (July 2003). "Wealth Accumulation over the Life Cycle and Precautionary Savings". In: *Journal of Business and Economic Statistics* 21(3), pp. 339–53.

Cagetti, M. and M. De Nardi (2006). "Entrepreneurship, Frictions and Wealth". In: *Journal of Political Economy* 114(5), pp. 835–70.

— (2009). "Estate Taxation, Entrepreneurship, and Wealth". In: *American Economic Review* 99(1), pp. 85–111.

Castañeda, A., J. Dıaz-Giménez, and J.-V. Rıos-Rull (Aug. 2003). "Accounting for the U.S. Earnings and Wealth Inequality". In: *Journal of Political Economy* 111, pp. 818–57.

Chang, Y. and S.-B. Kim (Dec. 2007). "Heterogeneity and Aggregation: Implications for Labor-Market Fluctuations". In: *American Economic Review* 97(5), pp. 1939–1956.

Cooley, T. F. and E. C. Prescott (1995). "Economic Growth and Business Cycles". In: *Frontiers of Business Cycle Research*. Ed. by T. F. Cooley. Princeton University Press, pp. 1–38.

De Nardi, M. (July 2004). "Wealth Inequality and Intergenerational Links". In: *Review of Economic Studies* 71, pp. 743–68.

De Nardi, M., G. Fella, and G. Paz-Pardo (2016). "The Implications of Richer Earnings Dynamics for Consumption and Wealth". In: *NBER working paper no. 21917*.

De Nardi, M., E. French, and J. B. Jones (2010). "Why do the Elderly Save?" In: *Journal of Political Economy* 118, pp. 39–75.

De Nardi, M. and F. Yang (2014). "Bequests and Heterogeneity in Retirement Wealth". In: *European Economic Review* 72, pp. 182–96.

Dynan, K., J. Skinner, and S. Zeld (2004). "Do the Rich Save More?" In: *Journal of Political Economy* 112, pp. 397–444.

Fagereng, A., L. Guiso, D. Malacrino, and L. Pistaferri (2020). "Heterogeneity and Persistence in Returns to Wealth". en. In: *Econometrica* 88(1), pp. 115–170. ISSN: 1468-0262.

Gale, W. G. and J. K. Scholz (1994). "Intergenerational Transfers and the Accumulation of Wealth". In: *Journal of Economic Perspectives* 8(4), pp. 145–60.

Gertler, M. (June 1999). "Government Debt and Social Security in a Life-Cycle Economy". In: *Carnegie-Rochester Conf. Ser. Public Policy* 50, pp. 61–110.

Glover, A. and J. Short (2015). "Bankruptcy, Incorporation, and the Nature of Entrepreneurial Risk". In: *Mimeo*.

Gokhale, J., L. J. Kotlikoff, J. Sefton, and M. Weale (2001). “Simulating the Transmission of Wealth Inequality via Bequests”. In: *Journal of Political Economy* 79(1), pp. 93–128.

Gollin, D. (Apr. 2002). “Getting Income Shares Right”. In: *Journal of Political Economy* 110, pp. 458–474.

Gourinchas, P.-O. and J. A. Parker (Jan. 2002). “Consumption over the Life Cycle”. In: *Econometrica* 70(1), pp. 47–89.

Gouveia, M. and R. Strauss (1994). “Effective Federal Individual Tax Functions: An Exploratory Empirical Analysis”. In: *National Tax Journal* 47(2), pp. 317–39.

Guvenen, F. (2009). “An Empirical Investigation of Labor Income Processes”. In: *Review of Economic Dynamics* 12, pp. 58–79.

Guvenen, F., G. Kambourov, B. Kuruscu, S. Ocampo, and D. Chen (2019). “Use It or Lose It: Efficiency Gains from Wealth Taxation”. en. In: *Working Paper*, p. 74.

Guvenen, F., F. Karahan, S. Ozkan, and J. Song (2016). “What Do Data on Millions of U.S. Workers Reveal about Life-Cycle Earnings Dynamics”. Mimeo.

Hendricks, L. (2004). “Accounting for Patterns of Wealth Inequality”. In: *Mimeo Iowa State University*.

— (2007). “How Important is Preference Heterogeneity for Wealth Inequality?” In: *Journal of Economic Dynamics and Control* 31(9), pp. 3042–68.

Huggett, M. (Dec. 1996). “Wealth Distribution in Life-Cycle Economies”. In: *Journal of Monetary Economics* 38, pp. 469–94.

Hurd, M. D. (July 1989). “Mortality Risk and Bequests”. In: *Econometrica* 57(4), pp. 779–813.

Jakobsen, K., K. Jakobsen, H. Kleven, and G. Zucman (Feb. 2020). “Wealth Taxation and Wealth Accumulation: Theory and Evidence From Denmark”. en. In: *The Quarterly Journal of Economics* 135(1), pp. 329–388. (Visited on 02/12/2020).

Kaldor, N. (Dec. 1957). “A Model of Economic Growth”. In: *Economic Journal* 67(268), pp. 591–624.

Kotlikoff, L. J., K. A. Smetters, and J. Walliser (July 1999). “Privatizing Social Security in the United States - Comparing the Options”. In: *Review of Economic Dynamics* 2, pp. 532–574.

Kotlikoff, L. J. and L. H. Summers (Aug. 1981). “The Role of Intergenerational Transfers in Aggregate Capital Accumulation”. In: *Journal of Political Economy* 89(4), pp. 706–32.

Krueger, D., K. Mitman, and F. Perri (2017). “On the Distribution of the Welfare Losses of Large Recessions”. In: *Handbook of Macroeconomics*.

Krusell, P. and A. A. Smith (Oct. 1998). “Income and Wealth Heterogeneity in the Macroeconomy”. In: *Journal of Political Economy* 106(5), pp. 867–896.

Kuhn, M. and J.-V. Rios-Rull (2016). “2013 update on the U.S. earnings, income, and wealth distributional facts : a view from macroeconomics”. English. In: *Federal Reserve Bank of Minneapolis Quarterly Review* 37(1), pp. 2–73.

Lawrence, E. (1991). “Poverty and the Rate of Time Preference: Evidence from Panel Data”. In: *Journal of Political Economy* 99, pp. 54–77.

Moll, B. (Oct. 2014). “Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation?” In: *American Economic Review* 104(10), pp. 3186–3221.

Paz-Pardo, G. (2016). “Preference Heterogeneity and Wealth Inequality”. In: *Mimo UCL*.

Quadrini, V. (Jan. 1999). “The Importance of Entrepreneurship for Wealth Concentration and Mobility”. In: *Review of Income and Wealth* 45, pp. 1–19.

— (Jan. 2000). “Entrepreneurship, Saving, and Social Mobility”. In: *Review of Economic Dynamics* 3, pp. 1–40.

Smith, M., D. Yagan, O. Zidar, and E. Zwick (Nov. 2019). “Capitalists in the Twenty-First Century*”. en. In: *The Quarterly Journal of Economics* 134(4), pp. 1675–1745. DOI: 10.1093/qje/qjz020.

Stokey, N. L. and S. Rebelo (June 1995). “Growth Effects of Flat-Rate Taxes”. In: *Journal of Political Economy* 103, pp. 519–550.

Storesletten, K., C. I. Telmer, and A. Yaron (June 2004). “Cyclical Dynamics in Idiosyncratic Labor Market Risk”. In: *Journal of Political Economy* 112, pp. 695–717.

Tauchen, G. and R. Hussey (Mar. 1991). “Quadrature-Based Methods for Obtaining Approximate Solutions to Nonlinear Asset Pricing Models”. In: *Econometrica* 59, pp. 371–396.

Zhang, H. H. (Dec. 1997). “Endogenous Borrowing Constraints with Incomplete Markets”. In: *Journal of Finance* 52(5), pp. 2187–2209.