

Saving and Wealth Inequality

Mariacristina De Nardi and Giulio Fella*

May 2017

Abstract

Why are some people wealth rich while others are poor? To what extent can governments affect inequality? Which instruments should they use? Answering these questions requires understanding why people save. Dynamic quantitative models of wealth inequality can help us to understand and quantify the determinants of the outcomes that we observe in the data and to evaluate the consequences of policy reform. This paper surveys the savings mechanisms generated by the transmission of bequests and human capital, by preference heterogeneity, by rate of return heterogeneity, by entrepreneurship, by richer earnings processes, and by medical expenses. It concludes that the transmission of bequests and human capital, entrepreneurship, and medical-expense risk are crucial determinants of savings and wealth inequality and that we need to look at more data to measure their relative importance.

Keywords: Human Capital; Bequests; Taxation; Entrepreneurship; Rates of Return; Earnings Shocks.

JEL Classification: E21; D14, D31.

*University College London, Federal Reserve Bank of Chicago, IFS, CEPR, and NBER, denardim@nber.org. Giulio Fella: Queen Mary University of London, CFM, and IFS, g.fella@qmul.ac.uk. A previous version of this paper by De Nardi was circulated under the title “Quantitative Models of Wealth Inequality: A Survey” and was the basis of her 2016 SED plenary lecture. De Nardi gratefully acknowledges support from the ERC, grant 614328 “Savings and Risks.” We thank Orazio Attanasio, Marco Bassetto, Richard Blundell, Matthias Doepke, Tim Lee, Cormac O’Dea, Vasia Panousi, José-Victor Ríos-Rull, and two anonymous referees for helpful comments. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research, any agency of the federal government, or the Federal Reserve Bank of Chicago.

1 Introduction

Why are some people wealth rich¹ while others are poor? To what extent can governments affect inequality? Which instruments should they use? Answering these questions requires understanding why people save. In fact, in many countries wealth is much more unequally distributed than labor earnings and income, and the wealthy keep saving at high rates. Dynamic quantitative models of wealth inequality can help us to understand and quantify the determinants of the wealth outcomes that we observe in the data and to evaluate the consequences of policy reform affecting them.

This survey starts from some basic facts about wealth inequality. It then introduces the workhorse framework for studying wealth inequality, the Bewley (1977) model, which features an incomplete market environment in which people save to self-insure against idiosyncratic earnings shocks. In this basic framework, precautionary savings in the face of earnings risk are the key force driving wealth concentration. However, since the ability to self-insure improves when wealth is large relative to earnings, the nature of precautionary savings implies that the saving rate decreases and then turns negative when one's net worth is large enough relative to one's labor earnings. Hence, the saving rate of the wealthy is negative in the basic Bewley model. In addition, the life-cycle version of the model also overestimates the fraction of people with little to no saving. Both of these implications are in contrast with the data. In the U.S. data for instance, rich people keep saving at high rates, which explains the emergence and persistence of very large fortunes, and the fraction of people with no savings at all is relatively small.

Moreover, the basic version of the Bewley model contains very few saving motives and might thus match the savings of some people for the wrong reasons. For instance, out-of-pocket medical expenses (including long-term-care) are likely an important reason to save for many households, including the upper-income ones. A model abstracting from this risk requires more patient agents to match observed net worth. As a consequence, the model might predict that people do not value government-provided health insurance because they are patient and do not face health risks.

The survey then discusses previous work that has uncovered forces that, first, have been shown to be empirically important and that, second, when introduced into a Bewley model, help improve the fit of the wealth distribution. These forces include the transmission of bequests and human capital across generations, heterogeneity in preferences, richer earnings processes, medical expense risk, heterogeneity in rates of return, and entrepreneurship.

The first force, the intergenerational transmission of bequests and human capital, is

¹In the interest of readability, we use “rich” instead of “wealth rich” in what follows, unless further qualification is called for.

large in the aggregate economy. Hence a natural question is whether it also has important implications for the distribution of wealth, in addition to its total amount. It turns out that introducing voluntary bequests of the luxury-good type and transmission of earnings ability between parents and children in a Bewley model generates more wealth concentration at the top, because some wealth is willingly transmitted across generations by the richer households, who in turn also tend to be higher earners. But, it also happens that, when calibrated using a standard earning process, this economy generates too many poor people.

The second force, heterogenous preferences, has been extensively documented in the empirical literature using a variety of methodologies. This mechanism, however, when introduced in a Bewley model, has had limited success in generating realistic inequality through the whole wealth distribution. This holds especially true in a life-cycle framework, unless one also realistically models bequest motives and the transmission of human capital. This mechanism can thus amplify other forces generating wealth inequality rather than being a crucial force driving the bulk of wealth inequality, especially at the top end.

The third force is earnings dynamics. Here, too, there is a vast and growing empirical literature documenting that earnings dynamics are much richer than typically assumed in these models. Typical assumptions are that, first, earnings follow a linear process in which the mean reversion and variance of earnings shocks do not depend on age or earnings levels; and second, that earning shocks are log-normally distributed and hence positive, and negative (log-)earnings shocks are equally likely. These assumptions are at odds with the data. This matters because sufficiently high negative skewness in earnings shocks can, in principle, generate large wealth concentration at the top. For these reasons, this survey turns to a Bewley model that includes an earnings process whose implications are consistent with the data along all of these dimensions. Its key finding is that richer modeling of earnings dynamics helps understand the saving decisions of the bottom 60% of the wealth distribution and helps explain the evolution of consumption inequality over the life cycle, but does not generate the kind of saving behavior at the top that is necessary to concentrate a lot of wealth in the hands of the richest.

The fourth set of forces that we discuss in this survey are medical expense risk and heterogeneity in life expectancy. Here, too, there is convincing evidence that the retirement period is one in which households face large income risk in the form of medical expenses and heterogeneous lifespans. In particular, the data show that out-of-pocket medical expenses (the portion of medical expenses that people end up paying) increase fast with age and lifetime income after age 80 and that people with higher lifetime income live significantly longer. Introducing out-of-pocket medical-expense risk and heterogenous longevity into a model of savings after retirement helps match wealth holdings by age and lifetime income quintiles during retirement and the lack of wealth decumulation that

is observed for the high-lifetime-income people even at advanced ages. These findings suggest that medical expenses after retirement are an important reason to save and that their role in generating savings and wealth inequality over all of the life cycle should be further studied.

The fifth force, idiosyncratic random shocks to the rate of return to wealth accumulation is a well-known theoretical mechanism capable of generating a long right tail in the wealth distribution. Heterogeneity in rates of return has been documented to be large empirically, including within asset classes, to be persistent over time, and to be correlated with entrepreneurial activity. Life-cycle models with rate of return shocks also require a luxury-good bequest motive to help generate the observed degree of wealth concentration. In fact, it has been found that luxury-bequest motives are quantitatively more important than heterogeneous rates of return. In addition, it is important to keep in mind that rates of return are endogenous to entrepreneurial and portfolio decisions and that their determinants should be studied.

The sixth and last force, entrepreneurship, is also supported by strong empirical evidence documenting that: (a) many entrepreneurs are rich and a large fraction of rich people are entrepreneurs; (b) entrepreneurs have a high saving rate both before and after entry; (c) entrepreneurs face some borrowing constraints. Entrepreneurship is an important way to endogenize rates of return by explicitly modeling their production function, borrowing constraints, and risks. The survey thus proceeds to analyze the role of entrepreneurship in the context of a Bewley model of inequality. It shows that in a model with a simple life-cycle structure, entrepreneurship not only generates realistic amount of wealth concentration and wealth mobility over time for both entrepreneurs and workers and a realistic fraction of entrepreneurs among the wealthiest, but that it also matches the role of entrepreneurs in hiring labor and employing capital. These findings thus indicate that entrepreneurial activity is an important force driving wealth concentration.

This survey finds that explicitly modeling the life cycle² is important and that the transmission of human capital and bequests across generations, entrepreneurship, and medical-expense risks are important determinants of saving and wealth inequality. It also argues that a rich model including these forces should be taken to additional data, in addition to wealth inequality, to disentangle their relative importance. These additional moments should include wealth mobility, both within and across generations, for both entrepreneurs and workers, portfolio composition, the marginal propensity to consume and bequeath out of an additional dollar, and the correlation between lifetime earnings and wealth at retirement.

This survey concludes by mentioning other important economic forces that so far have

²See also Krueger et al. (2016a).

not been examined in the context of quantitative models of wealth inequality and that are potentially important avenues for future research. They include modeling the family as a source of both risks (from both partners' wages and medical expenses and from divorce and death) and insurance (including joint savings and labor supply of two partners), endogenizing wages and human capital, endogenizing health, and endogenizing rates of return from assets. Finally, while we mainly have focused on the determinants of inequality at a point in time, much more work is needed to understand the dynamics of inequality and its determinants over time.

2 Some facts about wealth inequality

Key facts about the distribution of wealth have been highlighted in a large number of studies, including Wolff (1992, 1998), Davies and Shorrocks (2000), Kennickell (2003), Cagetti and De Nardi (2008), and Kuhn and Ríos-Rull (2016).

The most striking aspect of the wealth distribution in the United States is its degree of concentration. Over the past 30 years or so, households at the top 1% of the wealth distribution have held about one-third of the total wealth in the economy, and those in the top 5% have held more than half. At the other extreme, many households (more than 10%) hold little in assets. While there is agreement that the share held by the richest few is high, the extent to which it has changed over time (and why) is still subject to debate (Piketty (2014), Saez and Zucman (2014), Bricker et al. (2015), and Kopczuk (2014)).

An important related observation is that the concentration of wealth is much higher than that of earnings and income (Díaz-Giménez et al., 1997; Rodriguez et al., 2002). For example, in 1992 the Gini indexes for labor earnings, income (inclusive of transfers), and wealth were, respectively, .63, .57, and .78 (Díaz-Giménez et al., 1997), while in 1995 they were .61, .55 and .80 (Rodriguez et al., 2002).

In addition, the correlation between labor earnings, income, and wealth is positive, but well below one. Consistent with these findings, Hendricks (2007a) finds that the correlation coefficient between lifetime earnings and wealth at retirement (0.61) is much less than unity.

Several studies have documented significant differences in saving behavior across various groups that might help shed light on the above facts. (See Browning and Lusardi (1996) for a review of the literature.) In particular, Dynan et al. (2004) find a strong positive association between lifetime income and saving rates in U.S. data. De Nardi et al. (2010) show that, among the elderly, people with higher lifetime income not only reach retirement with more wealth, but also run down their net worth more slowly during the retirement period. They also show that the patterns of out-of-pocket medical spending help to account for the high wealth holdings of higher-income people during

retirement. Quadrini (1999) documents that entrepreneurs, who tend to be among the richest households, exhibit higher saving rates. Buera (2006, 2009) finds high saving behavior for entrepreneurs, both before and after entering entrepreneurship, thus indicating that people might save to both enter and expand their business.

Beyond cross-sectional inequality at a point in time, the degree of mobility within the earnings and wealth distributions—the extent to which rich households stay rich and poor households stay poor—is an additional important dimension. Hurst et al. (1998) use the Panel Study of Income Dynamics (PSID) to analyze wealth mobility between 1984 and 1994 and document that most of the mobility occurs in the mid-range deciles, while the top and bottom ones show high persistence. Using the same dataset, Quadrini (1999) studies the wealth mobility for entrepreneurs and non-entrepreneurs and finds that entrepreneurs are more upwardly mobile. Unfortunately, top-coding in the PSID does not provide a very accurate picture of what happens in the top percentiles. Progress has been made by Guvenen et al. (2015a) by analyzing administrative tax data for earnings in the U.S.

These facts not only help inform about potential saving motives, but also help discipline their strength and dynamics over time. At least a subset of these facts will be used in turn, together with other facts, to discipline each of the quantitative models that we now analyze.

3 Basic Bewley models, saving behavior, and wealth inequality

Bewley models are incomplete-market models in which households are usually ex-ante identical,³ in the sense that they face the same stochastic process for an endowment shock, but are ex-post heterogeneous, because they receive different sequences of shock realizations. An exogenously specified earnings process is typically the source of these shocks, and its properties are usually estimated from micro-level data on earnings. Aiyagari (1994) and Hansen and İmrohorođlu (1992) provide early general-equilibrium versions of Bewley models.

Formally, the canonical Bewley model features a population of agents maximizing expected utility over the remaining, possibly infinite, lifetime ($T - h$) subject to a multi-

³See Ljungqvist and Sargent (2000) for an overview of Bewley models (sometimes also called Aiyagari-Bewley-Huggett-İmrohorođlu models), including properties and solution methods. See Quadrini and Ríos-Rull (1997) for a discussion of why we need incomplete-market models to study wealth inequality.

period budget constraint; namely

$$\max_{\{c_t, a_{t+1}\}_{t=h}^T} \mathbb{E} \sum_{t=h}^T S_{t,h} \beta^{(t-h)} \frac{c_t^{1-\sigma}}{1-\sigma}$$

$$a_{t+1} = (1+r)a_t + y_t - c_t, \quad a_{t+1} \geq \underline{a},$$

where c_t denotes consumption at age t , a_t the asset stock, y_t the realization of the stochastic labor earnings, and r the return to the single risk free asset. The probability that the household survives to period t is $S_{t,h} = \prod_{l=h}^{t-1} s_l$, where s_l is the survival probability between age $l-1$ and age l . In each period t , the household allocates total resources between current consumption and next period's assets, subject to the borrowing limit \underline{a} . Labor earnings are usually assumed to follow a first-order Markov process.

While computing the transitional equilibrium dynamics is sometimes feasible, these models are often solved for stationary equilibria. Since it is assumed that there is no aggregate uncertainty, in a stationary equilibrium there is a constant distribution of people over state variables. However, individuals face considerable uncertainty as they move up and down the distribution.

These models endogenously generate differences in asset holdings as a result of the household's desire to save and the realization of the exogenous shocks. Incomplete-market models can be applied to study many interesting and important questions that go beyond wealth inequality and, thus, the scope of this survey. See Quadrini and Ríos-Rull (2014), Krusell and Smith (2006), Guvenen (2016), and Heatcote et al. (2009) for surveys on this, and Conesa et al. (2009), Krueger and Fernandez-Villaverde (2011), and Krueger et al. (2016b) for interesting applications.

3.1 The infinitely-lived Bewley model

In the infinitely-lived case, the probability of survival s_t is identically equal to one and the stochastic process for earnings is time-independent. A version of this model is quantified by Aiyagari (1994), who assumes that (log) yearly labor earnings follow a first-order autoregressive process, with an autocorrelation of 0.6 and a standard deviation of innovations of 0.2. This results in an unconditional coefficient of variation of 0.31.⁴ Aiyagari also considers a process with twice the standard deviation of the innovation for earnings, which results in an unconditional coefficient of variation of 0.63; this is a much higher variability process than typically estimated in the literature. Quadrini and Ríos-Rull (1997) summarize the implications of the model for these two parameterizations of the earnings process.

⁴These figures are based on estimates from Abowd and Card (1989) using micro-level panel data.

Wealth Gini	% Wealth in top		
	1%	5%	20%
U.S. data, 1989 SCF			
.78	29	53	80
Aiyagari Baseline			
.38	3.2	12.2	41.0
Aiyagari higher variability			
.41	4.0	15.6	44.6

Table 1: A Bewley model with infinitely-lived agents. Data from the 1989 Survey of Consumer Finances (SCF) in the top line of data and corresponding simulated models in the bottom two lines of data, as reported by Quadrini and Ríos-Rull (1997).

Table 1 reports their findings for the wealth distribution. The first row refers to data from the 1989 Survey of Consumer Finances (SCF). The second and third rows report the corresponding moments for respectively the baseline calibration and the one with higher earnings volatility in Aiyagari (see Quadrini and Ríos-Rull (1997)). The comparison makes clear that this version of the model comes nowhere near to matching the Gini coefficient, let alone the degree of wealth concentration among the top 20% or less of individuals. For instance, the richest 1% of people in these versions of the model hold, at most, 4% of total net worth, compared with 29% in the data, and the Gini coefficient generated by the model is half the one in the data.

3.2 A basic overlapping-generation Bewley model

The infinite horizon Bewley model does not account for the heterogeneity that arises from the life cycle, which is an important source of heterogeneity in wealth because people typically enter the labor market with little to no assets and then gradually accumulate them, at least until retirement age.

To introduce a life-cycle dimension, assume that in each period a continuum of agents are born. Each agent lives at most T periods and faces an age-dependent survival probability s_t . Surviving agents work up to age $L < T$ and retire afterwards. The demographic patterns are assumed to be stable, hence age- t agents make up for a constant fraction μ_t of the population at every point in time.

The earnings process is now age-dependent. During the working period, it is composed of a deterministic component, which is hump-shaped by age, and a stochastic component, which is a first-order Markov chain. During retirement, it equals a constant Social Security benefit.

There are no annuity markets⁵ to insure against mortality risk. People who die pre-

⁵This is a commonly used assumption because the annuity market is small in practice. Eichenbaum

maturely leave accidental bequests, which are redistributed uniformly among all people alive. Compared with the previous framework with infinitely-lived agents, two more saving motives are present: to smooth consumption during retirement and to self-insure against longevity risk.⁶ In principle, these additional saving motives could generate more wealth inequality and higher saving rates than in the model with infinite lifetimes.

Huggett (1996) calibrates this model economy to key features of U.S. data and uses different versions of it to quantify how much wealth inequality it can generate.

Transfer wealth ratio	Wealth Gini	Percentage wealth in the top					Percentage with negative or zero wealth
		1%	5%	20%	40%	60%	
1989 U.S. data							
.60	.78	29	53	80	93	98	6
A basic overlapping-generations Bewley model							
.67	.67	7	27	69	90	98	17

Table 2: A basic overlapping-generations Bewley model, from De Nardi (2004b).

Table 2 compares wealth moments from the data and a calibrated version of Huggett’s model. The first row refers to the 1989 U.S. data. The second one refers to De Nardi’s (2004b) version of Huggett’s model⁷ with only accidental bequests. Compared with the infinite-horizon model, Huggett’s model economy succeeds in matching the U.S. Gini coefficient for wealth, but the concentration is obtained by having too many people holding little wealth, rather than by concentrating enough wealth in the right tail of the wealth distribution.

The key reason for this failure is that in the data, the rich have a high saving rate, while in the model, households stop saving once they have accumulated a sufficiently high buffer-stock (Carroll, 1997) and retirement saving. This is illustrated in Figure 1, which reports the saving rate as a function of current wealth for an individual with median earnings (corresponding to \$32,000 in the model, expressed in year 2000 \$) at different ages. Even for individuals close to retirement, the target level of wealth does not exceed ten times current earnings. Thus, the additional saving motives in this version of the model (saving for retirement and for longevity risk) help bring the implications of the

and Peled (1987) show that in the presence of moral hazard, people will choose to self-insure rather than use annuity markets, even if the rate of return on annuities is high.

⁶For the same reason, modeling Social Security explicitly is important because Social Security redistributes a significant fraction of income from the young to the old and, thus, reduces the saving rate and changes the aggregate capital-output ratio.

⁷These results are very similar to Huggett’s, though they refer to a model period of five years. We report this version for easier comparability with the results on the transmission of bequests and human capital. The length of the time period is the main reason why these results are slightly different from those in the benchmark in De Nardi et al. (2016a), which uses a one-year model period.

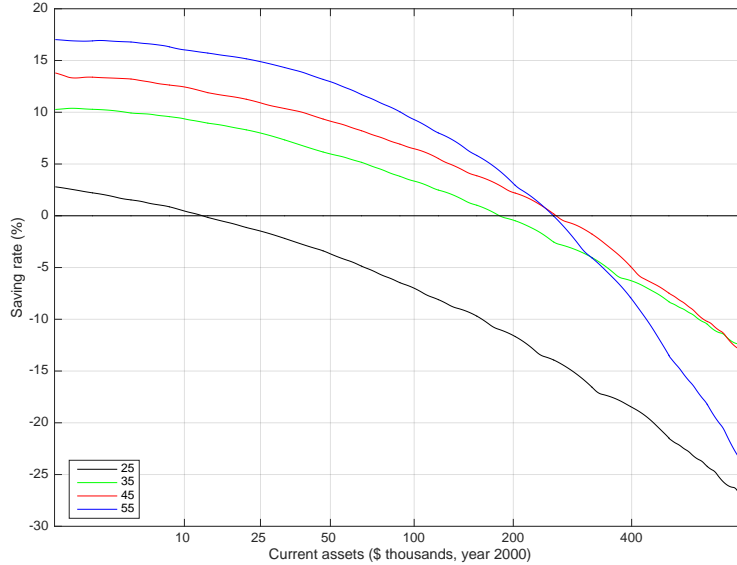


Figure 1: Saving rate by age and current wealth for an individual with median earnings (corresponding to \$32,000 in the model, expressed in year 2000 \$).

model closer to the data, but do not go far enough in that direction, as they do not sufficiently raise the saving rate of people as they get richer.

Huggett also finds that relaxing the household’s borrowing constraint increases the fraction of people bunched at zero or negative wealth, but does not increase much the asset holdings of the rich. Hence, it does not help in generating a distribution of wealth closer to the observed one. In addition, he documents the amount of wealth inequality generated by his model at different ages and shows that, starting from age 40, the model underpredicts the amount of wealth inequality.

To sum up, this class of models is not capable of explaining why rich people keep saving at high rates. In particular, it implies that, to the extent that earnings follow a canonical (log-) linear process (e.g. Abowd and Card, 1989), the right tail of the wealth distribution cannot be thicker than that of the distribution of earnings. As discussed above, this is at odds with the empirical evidence that wealth is substantially more concentrated than earnings.

3.3 Lessons learned from the basic quantitative Bewley model

The two previous subsections thus show that both the infinite-horizon and the life-cycle versions of the Bewley model do not match well the observed distribution of wealth. In particular, while the life-cycle version improves the fit of the wealth Gini coefficient, it does so by generating rich people who are not nearly rich enough, middle-class people who are too rich, and poor people who are too poor compared with the actual data.

A number of empirically important economic forces have been included in the basic

Bewley setup to assess the extent to which they improve its ability to match the observed wealth inequality. Because saving behavior depends on preferences and the accumulation technology implied by the dynamics budget constraint (more specifically, non-capital earnings risk and the rate of return on wealth), we classify the economic forces considered into these two components of the model in what follows.

4 Bequests and transmission of human capital

Intergenerational transmission of wealth is empirically important. Kotlikoff and Summers (1981) argue that it accounts for the majority of aggregate capital formation. Further studies have found that intergenerational transfers account for at least 50-60% of total wealth accumulation (Gale and Scholz (1994))⁸ Given that intergenerational transfers are large in the aggregate, they might also play an important role in shaping wealth inequality.

On the theory side, Becker and Tomes (1979) were the first to model the parental decision problem and to characterize the transmission of both human capital and bequests across generations. They showed that in the presence of borrowing constraints, parental transfers first come in the form of children’s human capital investment; and that only after the optimal amount of human capital investment in children has been achieved, parents find it optimal to start giving monetary transfers, such as bequests. Bequests are thus a luxury good in this framework.

Further developing these ideas in a quantitative framework, De Nardi (2004b) introduces two types of intergenerational links in the OLG model used by Huggett: voluntary bequests and transmission of human capital. She models the utility from bequests as providing a “warm glow” (as in Andreoni (1989)) and the transmission of human capital as the correlation between children’s labor earnings at labor market entry with parent’s labor earnings at the same time. In this framework, parents and their children are thus linked by voluntary and accidental bequests and by the transmission of earnings ability. The households thus save to self-insure against labor earnings shocks and life-span risk, for retirement, and possibly to leave bequests to their children. Thus, this version of the model changes both preferences and technology (and more specifically the endowment).

More specifically, compared with the standard Bewley model, the voluntary “warm glow” bequest motive introduces an extra utility term. The individual’s problem is therefore

⁸Wolff and Gittleman (2014) and Kuhn and Ríos-Rull (2016) suggest somewhat smaller numbers, but it is well known that the exact way that this number is computed is important. See Davies and Shorrocks (2000).

$$\max_{\{c_t, a_{t+1}\}_{t=h}^T} \mathbb{E} \sum_{t=h}^T S_{t,h} \beta^{(t-h)} \left(\frac{c_t^{1-\sigma}}{1-\sigma} + \boxed{(1-s_{t+1})\phi(b(a_{t+1}))} \right)$$

$$a_{t+1} = (1+r)a_t + y_t - c_t, \quad a_{t+1} \geq \underline{a}$$

where the term

$$\phi(b(a_{t+1})) = \phi_1 \left(1 + \frac{b(a_{t+1})}{\phi_2} \right)^{1-\sigma}$$

captures the utility of bequeathing $b(a_{t+1})$, where the function $b(\cdot)$ maps the parent's wealth at death into the bequest, net of estate taxes, received by the offspring. The utility from leaving bequests depends on two parameters: ϕ_1 , which represents the strength of the bequest motive, and ϕ_2 , which measures the extent to which bequests are a luxury good because it affects the marginal utility of bequests in a nonlinear way.⁹ These two parameters are respectively calibrated to data on the fraction of capital due to intergenerational transfers and the 30% share of singles that leave estates of little or no value.

It should be noted that some papers that do not find evidence in favor of a bequest motive (e.g. Hurd, 1989; Hendricks, 2004) assume that utility is homothetic in bequests ($\phi_2 = 0$), thus generating the counterfactual implication that even poor people save to leave bequests of significant size. In addition, Hurd (1989) identifies a low marginal propensity to leave bequests by comparing the asset trajectories of households with and without children, and we now know that people without children also want to leave bequests (Kopczuk and Lupton (2007)). In contrast, looking at a sample of wealthier retirees, Laitner and Juster (1996) find that about half of the households in their sample plan to leave estates and that the amount of wealth attributable to estate building is significant, accounting for half or more of the total for those who plan to leave bequests.

De Nardi's flexible functional form and parameterization imply a realistic distribution of estates. Her calibration is also quantitatively consistent with the estimates of the elasticity of the savings of the elderly to permanent income by Altonji and Villanueva (2003).

Table 3 summarizes De Nardi's main results. The first two rows, reported for convenience, are the same as in Table 2 and refer, respectively, to the 1989 SCF U.S. data and the version of Huggett's model economy with only accidental bequests, redistributed uniformly in every period. The third row refers to an economy in which there are only accidental bequests that, rather than being redistributed uniformly, are received by the children of the deceased only once, upon their parent's death. This formulation implies

⁹See De Nardi (2004b) for more discussion on this.

Transfer wealth ratio	Wealth Gini	Percentage wealth in the top					Percentage with negative or zero wealth
		1%	5%	20%	40%	60%	
1989 U.S. data							
.60	.78	29	53	80	93	98	6
No intergenerational links, equal bequests to all							
.67	.67	7	27	69	90	98	17
No intergenerational links, unequal bequests to children							
.38	.68	7	27	69	91	99	17
One link: parent's bequest motive							
.55	.74	14	37	76	95	100	19
Both links: parent's bequest motive and productivity inheritance							
.60	.76	18	42	79	95	100	19

Table 3: OLG models of wealth inequality, from De Nardi (2004b).

that bequests are both unequally distributed and received at a realistic age, rather than every period).

Comparing rows two and three reveals that allowing for a more realistic timing and distribution of accidental bequests does not generate a more unequal wealth distribution. This is because receipt of a bequest per se does not alter the saving behavior of the richest. On the other hand, the timing of transfers does significantly affect the transfer-wealth ratio in the first column of the table. The transfer-wealth ratio—the ratio of wealth transmitted across generations to aggregate capital—is a measure of intergenerational transfers first proposed by Kotlikoff and Summers (1981). The ratio is sensitive to the timing of transfers because of the way that transfers are capitalized and accumulated interest accrues to bequests. If children inherit only once, when their parent dies (rather than every year), then the fraction of wealth attributed to intergenerational transfers in the model is much lower than the one in the data.

The fourth row in Table 3 allows for a voluntary bequest motive and shows that voluntary bequests can explain the emergence of large estates, which are often accumulated in more than one generation and are important for the upper tail of the wealth distribution in the data. The bequest motive to save is much stronger for the richest households, who, even when very old, keep some assets to leave to their children. The rich leave more wealth to their offspring, who, in turn, tend to do the same. In steady state, this behavior generates some large estates that are transmitted across generations because of the voluntary bequests.

The fifth row allows for both voluntary bequests and transmission of ability and shows that a human-capital link through which children partially inherit the productivity of their parents generates an even more concentrated wealth distribution. More productive

parents accumulate larger estates and leave larger bequests to their children, who, in turn, are more productive than average in the workplace.

Therefore, a luxury-type bequest motive can help to explain why rich households save at much higher rates than the rest (Dynan et al., 2004).¹⁰ As shown in Figure 2, the presence of a luxury-type bequest motive also generates lifetime saving profiles that imply slower wealth decumulation in old age for richer people, consistent with the facts documented by De Nardi et al. (2010), using micro-level data from the Health and Retirement Survey.

In a model with intergenerational links that abstracts from medical-expense risk, saving for precautionary purposes and saving for retirement are the primary factors for wealth accumulation at the lower tail of the distribution, while saving to leave bequests significantly affects the shape of the upper tail.

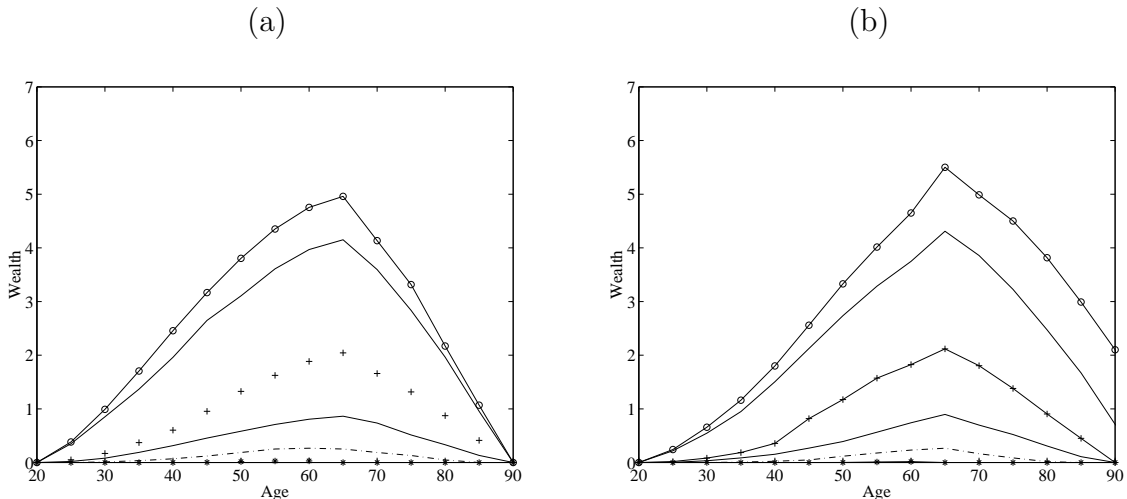


Figure 2: Age profile of average wealth for .1, .3, .5, .7, .9, .95 quantiles. No links, equal bequests to all, panel (a), and bequest motive, panel (b).

This model also has implications for wealth heterogeneity at retirement. Venti and Wise (1998) and Bernheim et al. (2001) show that wealth is highly dispersed at retirement, even for people with similar lifetime incomes. Bernheim et al. (2001) argue that the observed differences are hard to explain in the context of a model with rational agents and are, in contrast, better explained by “rule of thumb” or “mental accounting” behavior. A few papers further investigate the implications of models with rational agents along these dimensions. Hendricks (2004) studies the implications of a basic OLG model with accidental bequests and shows that, at retirement age, the model overstates wealth differences between earnings-rich and earnings-poor households, while it understates the amount of wealth inequality conditional on lifetime earnings. In contrast, De Nardi and Yang (2014)

¹⁰An, effectively isomorphic, mechanism is Carroll’s (2000) “capitalist spirit” model, in which finitely-lived consumers have wealth in the utility function, which can be calibrated to make wealth a luxury good, thus generating nonhomothetic preferences.

show that, when augmented with voluntary bequests and intergenerational transmission of earnings, the OLG model can match the observed cross-sectional differences in wealth at retirement and their correlation with lifetime incomes.

Gokhale et al. (2001) also account for wealth inequality at retirement in an overlapping-generations model with infinite risk aversion, only accidental bequests, and a rich set of exogenous features (death and fertility, assortative mating, and heterogeneous human capital and rate of return). In their environment, skill differences, assortative mating, Social Security, and time preferences are the primary determinants of wealth inequality at retirement.

4.1 Interesting potential extensions

The papers by De Nardi (2004b) and De Nardi and Yang (2014) thus highlight some important mechanisms that are not only empirically relevant, but also help reduce the gap between the implications of the standard life-cycle Bewley model and observed wealth inequality. However, the framework that they adopt has a number of limitations and could thus be extended further.

First, it takes the transmission of human capital, or individual productivity, as exogenous. There is a vast literature on the endogenous transmission of human capital, but not on its implications for inequality in wealth holdings. For instance, Aiyagari et al. (2002) study optimal parental investment of time and money in children, both with perfect and imperfect altruism. Brown et al. (2011) develop a model in which parents and children make heterogeneous investments in children's education, and some parents underinvest in it. Lee et al. (2014) study the importance of parental investment on the intergenerational transmission of economic status, while Lee and Seshadri (2014) attempt to identify the causal effect of parental human capital on children's human capital. For surveys about the importance of parental background, see Heckman and Mosso (2014) and Bowles et al. (2009).

Second, the framework assumes that fertility is exogenous and that everyone has the same number of children. Scholz and Seshadri (2007) examine the effects of children in a life-cycle model with endogenous fertility. They argue that children have a large effect on household's net worth and consequently are an important factor in understanding the wealth distribution. They also find that fertility and credit constraints interact in ways that significantly affect wealth accumulation.

Third, there are no inter-vivos transfers between parents and children. Nishiyama (2002) adopts an OLG model with bequests and inter-vivos transfers in which households in the same family line behave strategically. Like De Nardi, he concludes that the model with inter-vivos transfers helps explain some of the large fortunes that are observed in

the data, thus confirming that transfers across generations before and after death have similar implications in terms of wealth inequality.

Fourth, one might think that if households' voluntary bequest motives are an important reason why rich households keep saving, the specific bequest formulation might be quite important in determining the response to taxation. Interestingly, De Nardi and Yang (2015) find that, regardless of whether warm-glow bequests of the type that we have discussed in this paper depend on estates net or gross of taxes, the model does not generate very different responses to estate taxation reform as long as the models are calibrated to match the same facts. More investigation on the robustness of this result to different policies and formulations of the bequest motives is called for.

5 Preference heterogeneity

There is enough micro-level empirical evidence of heterogeneity in time preferences (Lawrence, 1991) and both time preferences and risk aversion (Cagetti, 2003) to suggest that preference heterogeneity might be a plausible avenue to help explain the vastly different amounts of wealth held by households in the data.

Introducing heterogeneity in patience and risk aversion in the standard Bewley model implies a modified utility function

$$\max_{\{c_t, a_{t+1}\}_{t=h}^T} \mathbb{E} \sum_{t=h}^T S_{t,h} \beta_i^{(t-h)} \frac{c_t^{1-\sigma_i}}{1-\sigma_i},$$

in which both the discount factor β_i and the risk-aversion coefficient σ_i are specific to individual i and may evolve stochastically.

Krusell and Smith (1998) extend the infinitely-lived version of the Bewley model by introducing a stochastic process for each dynasty's discount factor, implying it changes on average every generation. They keep, instead, risk aversion constant and homogeneous across dynasties. They show that a small degree of stochastic heterogeneity in discount factors allows the model to match the variance of the cross-sectional distribution of wealth and generates more wealth concentration among the richest. However, while capturing the variance of the wealth distribution well, their model and calibration fall short of matching the large wealth concentration at the top 1% (24% in the model, compared with 28-33% in the data, depending on the reference year).

Hendricks (2007b) studies the effects of stochastically evolving, discount factor heterogeneity in a life-cycle, rather than dynastic, framework with purely accidental bequests. To discipline preferences, he requires that his model matches the observed age pattern of the wealth Gini coefficient that he estimates from the data. After matching these

calibration targets, he examines the model's implications for the cross-sectional distribution of wealth. Contrary to the infinite-horizon setup of Krusell and Smith (1998), he finds time-preference heterogeneity makes only a modest contribution to accounting for the observed high wealth concentration. This result obtains despite the fact that, in his life-cycle model, the degree of heterogeneity in discount factors chosen to match the age-profile of the wealth Gini coefficients is much higher than in Krusell and Smith (1998). In addition, the implied gap between the discount factors of the most and least patient households is more than twice the corresponding gap estimated by Lawrence (1991), using consumption Euler equations across permanent income levels.

Hendricks (2007b) argues that the difference in results between the infinitely- and finitely-lived formulations is due to the fact that the presence of an additional (retirement) saving motive implies that, for the same parameter values, the life-cycle model implies a higher wealth-income ratio than the dynastic model. Therefore, if one denotes by $\beta_i(1+r)$ the effective degree of impatience, its average value consistent with a standard wealth-income ratio target of 2.5-3 is close to one in a dynastic model, but substantially below one in a life-cycle model. As a result, for any sensible degree of discount rate heterogeneity, the impatience condition $-\beta_i(1+r) < 1$ – guaranteeing a finite target for the ratio of wealth to permanent income holds for *all* agents in a life-cycle model. And, vice versa, in a dynastic model, the same degree of heterogeneity, combined with a substantially higher average, implies that $\beta_i(1+r) > 1$ for the most patient individuals. The violation of the impatience conditions (or equivalently the fact that they have an infinite wealth-income target) explains why the most patient individuals accumulate large amounts of wealth in the dynastic model.

Recent work by Paz-Pardo (2016) confirms Hendrick's findings in the context of a standard life-cycle model, but also shows that, in combination with a luxury-type bequest motive and productivity inheritance as in De Nardi (2004b), an empirically reasonable level of preference heterogeneity can help explain wealth concentration among the top 1% of non-business owners, while still generating a life-cycle increase in average consumption consistent with the data.

Heer (1999) adopts a model in which richer and poorer people have different tastes for leaving bequests to generate heterogeneity in wealth holdings, while Laitner (2001) assumes that all households save for life-cycle purposes but that only some of them are altruistic toward their children. Laitner allows for perfect annuity markets, therefore all bequests are voluntary, and there is no earning risk over the life cycle, hence no precautionary savings. In addition, he matches the concentration in the upper tail of the wealth distribution by choosing both the fraction of altruistic households and the distribution of wealth within the dynasty (which is indeterminate in the model).

More in the spirit of experimenting with the formulation of preferences, rather than of

allowing for preference heterogeneity, Díaz et al. (2002) study the effect of habit formation and find that it actually decreases the concentration of wealth generated by this type of model. In fact, habits act similarly to increased risk aversion, and more risk aversion tends to increase the saving of everyone and to dampen wealth dispersion.

5.1 Interesting potential extensions

In sum, previous work indicates that preference heterogeneity, and especially patience heterogeneity, can generate increased heterogeneity in wealth holdings to some extent. However, it is difficult to determine how much preference heterogeneity it is reasonable to incorporate in a model and whether we are overstating the role of this factor. More generally, it would be interesting to deepen the previous analysis by allowing for richer formulations of the utility function in which, for instance, risk aversion and intertemporal substitution do not have to coincide (see Wang et al. (2015), for some interesting findings on this) and for merging this source of inequality with other explanations generating inequality in wealth holdings, to better evaluate its importance in conjunction with others.

6 Earnings risk

The modeling of the earnings process has to do with the kind of technology (the endowment, more specifically) that is assumed in a Bewley model. The standard assumption in quantitative models of wealth inequality is that (log) labor earnings follow an age-independent linear process with homoskedastic Gaussian innovations around a deterministic age-efficiency profile. It should be noted that the age-independent linearity assumption implies that the mean reversion of a shock is constant by earnings levels and age.

Yet, Arellano et al. (2017) find that the persistence of earnings innovations depends both on age and previous earnings level; Geweke and Keane (2000) and Bonhomme and Robin (2009) document that innovations to earnings are not Gaussian; Meghir and Pistaferri (2004) and Blundell et al. (2015) show that innovations to earnings are not homoskedastic; and Guvenen et al. (2015b) highlight that earnings changes display substantial negative skewness and kurtosis and that the conditional moments of earnings changes display substantial variation by age and previous earnings level.¹¹

Castañeda et al. (2003) were the first to highlight how a stochastic process featuring negative skewness may help generate a long right tail in the wealth distribution. They consider a model economy with a two-stage life cycle (working time and retirement),

¹¹Guvenen et al. (2015b) document these features using U.S. Social Security Administration tax earnings (W2) data, while Arellano et al. (2017) use Norwegian tax data, but also show that similar features hold in the PSID.

Earnings level	1.0	3.0	10.0	1060
Fraction at invariant distribution	61.11%	22.25%	16.50%	0.04%

Table 4: Castañeda et al.’s (2003) earnings process.

in which workers have a constant probability of retiring in each period, and retirees face a constant probability of dying. Each household is perfectly altruistic toward its descendants.

The key feature of their model that generates a large amount of wealth holdings in the hands of the richest is the productivity shock process, whose key features are reported in Table 4. This process is calibrated to match features of both the earnings *and* wealth distributions in the United States. The calibration implies that the highest productivity level is more than 100 times higher than the second highest. Thus, there is a large discrepancy between the highest productivity level and all of the others. Moreover, an individual with the highest productivity level has a more than 20% chance of being 100 times less productive during the following period. High-earning households thus face higher earnings risk and save at high rates to self-insure against this risk. As a result they accumulate a large buffer stock of assets. This finding implies that an earnings process displaying sufficiently large negative skewness at the top end is capable of generating a long right tail in the wealth distribution.

Importantly, the properties of this earnings process are calibrated to match cross-sectional moments of the earnings *and* wealth distributions, rather than using household-level data on earnings dynamics over time. This was a forced choice at the time the paper was written because the panel data sets available at the time (e.g., the PSID) were not representative of earnings risk for the richest individuals, due to top coding and lack of oversampling at the top. The recent availability of large panel data sets that do not suffer from these shortcomings raises the question of whether the degree of negative skewness in actual earnings data is large enough to generate the observed levels of top wealth concentration.

De Nardi et al. (2016a) reappraise this question by studying the implications of a rich earnings process, consistent with the one estimated by Guvenen et al. (2015a), in an otherwise standard life-cycle model. More specifically, they compare the wealth distribution generated by the model in the case in which earnings follow: (1) the AR(1) process used by Huggett (1996); or, alternatively, (2) the rich earnings process implied by the estimates in Guvenen et al. (2015a).¹² Table 5 summarizes their findings. The first row refers to 1989 U.S. data from the SCF, while the second and third rows refer to the model

¹²In both cases, they calibrate the discount factor to match a target wealth-income ratio of 3 in the data.

Bequests- output ratio	Wealth Gini	Percentage wealth in the top						Percentage with negative or zero wealth
		1%	5%	20%	40%	60%	80%	
U.S. data 2.6%	.72	28	49	75	89	96	99	6
AR(1) - Huggett (1996) 2.8%	.72	12	35	74	93	99	100	14
Richer earnings process - De Nardi, Fella and Paz-Pardo (2016) 2.7%	.65	10	30	67	88	96	99.6	5

Table 5: Wealth distribution statistics under alternative earnings processes. Source: De Nardi et al. (2016a).

with, respectively, Huggett’s (1996) AR(1) and the richer earnings process.

Allowing for richer earnings dynamics improves the ability of the model to match the wealth holdings of the poorest 60% of the population. In particular, allowing for richer earnings dynamics dramatically improves the ability of the model to match the proportion of individuals with zero or little wealth. De Nardi et al. (2016a) show that the main driver for this result is that the richer earnings process implies conditional moments that change with the previous earnings realization. In particular, the persistence of low earnings realizations is substantially below the average degree persistence. As a result, the richer earnings process implies that low-earning individuals engage in more precautionary saving relative to the AR(1) process.

Turning to the right tail of the wealth distribution, though, despite the fact that the richer earnings process matches the degree of negative skewness in the data, this richer earnings process implies a degree of top wealth concentration similar to that implied by the AR(1). That is, the richest 1% hold only about 10% of total wealth, compared with at least 28% in the data. This finding suggests that, even when considering tax data, which should be representative also of individuals at the top of the earnings distribution, the degree of skewness in the earnings data is not sufficient to generate enough precautionary saving by the rich to match the observed wealth at the top.

An important caveat, however, is that the tax data used by Guvenen et al. (2015a) do not contain business income in privately held businesses and thus might not capture the income of business owners/entrepreneurs, which account for 60% of individuals in the top 1% of the wealth distribution. DeBacker et al. (2012) use a confidential panel of U.S. income tax returns for 1987-2009 to measure business income risks and document that, compared with labor income, business income is much riskier (even conditional on staying in business), less persistent over time, and characterized by higher probabilities of extreme upward or downward mobility. They also show that high-income entrepreneurs are more likely to face tail risk at both ends of the business income distribution. Both

of these findings are generally consistent with the idea that high earners are subject to larger fluctuations. Parker and Vissing-Jørgensen (2009) also find that incomes at the top are cyclical because of the labor component and bonuses, in particular. Although for business owners the split between their wages and capital income might be somewhat flexible, these authors write “High-income households (top 1 percent) earn more than half of their noncapital gains income from wage income, and their wage income is far more exposed to aggregate fluctuations than that of lower-income households...we find even higher income exposure to aggregate fluctuations for high-income households (top 0.01 percent) than for low-income households...”.

Interestingly, Barnett and Panousi (2015) also uncover that the risk faced by business people is heteroskedastic: high-wealth agents are more likely than low-wealth agents to have big business-income fluctuations (both big increases and big declines). In contrast, these “risks” do not vary along other dimensions, such as gender, level of education, and race.

One more caveat is in order. A substantial fraction of business income likely constitutes remuneration for the entrepreneur’s skill (labor earnings from the perspective of the model), rather than return on the owner’s capital investment (capital return). But in the absence of an appropriate criterion to apportion a share of business income to labor earnings, it is not clear how the properties of business owners’ income measured by DeBacker et al. (2012) should inform the calibration of labor earnings processes.

6.1 Interesting potential extensions

De Nardi et al. (2016a) show that earnings data for non-entrepreneurs do not feature sufficient downward risk to generate a long right-tail in the wealth distribution as a result of precautionary saving. To the extent that this kind of risk is confined to entrepreneurs or business owners, it should not just be modeled as an exogenous shock, but should, rather, be endogenous to entrepreneurial decisions about savings, labor hiring, and portfolio choice. It would be interesting to use the new and more detailed data on the risks faced by entrepreneurs to formalize and estimate a model that matches the earnings, and rate of return dynamics and heterogeneity, associated with entrepreneurial activity, and study its implications for wealth inequality among both entrepreneurs and the whole population.

7 Medical expenses

De Nardi et al. (2010) use the Health and Retirement Survey (HRS) data and find that after retirement, out-of-pocket medical, including nursing home, costs rise with age and permanent income (PI). Figure 3 shows that, especially after age 80, out-of-pocket medical

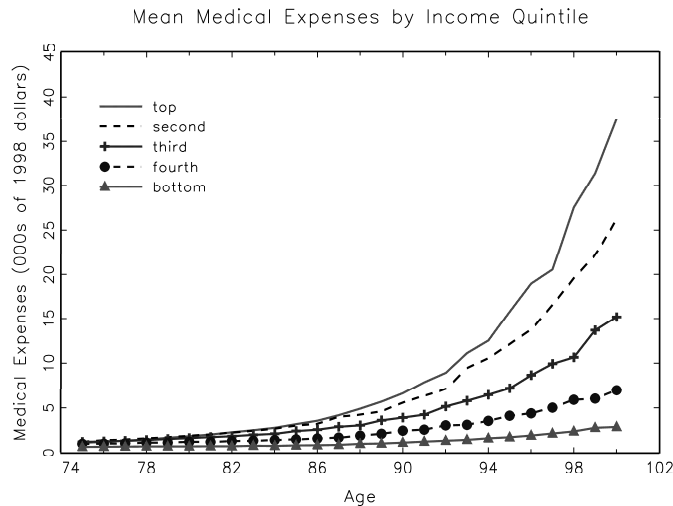


Figure 3: Average out-of-pocket medical expenses after age 74, by permanent income. HRS data

expenses increase fast by age for people in the highest PI quintile and, in particular, can surpass \$20,000 a year after age 95.

They also find that, during retirement, the elderly with the high PI dissave little until very advanced ages, the low PI elderly never save, while the middle PI elderly do dissave. More specifically, Figure 4 displays median assets, conditional on birth cohort and permanent income quintile, for singles (who tend to have lower assets than couples). It presents asset profiles for the unbalanced panel, each point displaying the median for all the members of a particular cell who are alive at a particular date. Median assets are increasing in permanent income, with the 74-year-olds in the highest PI quintile holding median assets of about \$200,000 and those in the lowest PI quintiles holding essentially no assets at all. Over time, those with the highest PI tend to hold onto significant wealth well into their nineties, those with the lower PIs never save much, while those in the middle PIs display some asset decumulation as they age.

They build a model of savings and medical expenses after retirement in which, consistent with the data, people with higher PIs also have longer life expectancies. Thus, compared with our Bewley models, preferences are heterogeneous because survival probabilities are heterogeneous by PI, hence, the effective discount factor, given by the product of β and s_i^t are heterogeneous, where s_i^t indexes individual heterogeneity in survival probabilities. In addition, there is also a modification to the basic model's technology. More specifically, medical expenses hit the budget constraint (as in Hubbard et al. (1995) and Hubbard et al. (1994)) as an exogenous shock to resources. The budget constraint is thus modified as follows

$$a_{t+1} = (1 + r)a_t + y_t - m_t, \quad a \geq \underline{a},$$



Figure 4: Savings after age 74, by age and permanent income (HRS data).

where m_t is the out-of-pocket medical expenses shock. In addition, the government provides a consumption floor.

Figure 5 shows that the model with medical expenses fits the savings data after age 74, by age and PI well. Medical expenses that increase with age and permanent income are an important reason why the high PI elderly do not run down their assets, while government insurance covers the low PI households, who never save during retirement. Thus, based on their work, medical expenses (including for long-term care) and government insurance programs have large and heterogenous effects on savings.

Kopecky and Koreshkova (2014) separately allow for medical and nursing-home expenditure risk during retirement and consider their effect on saving over the whole life cycle rather than only after retirement. They find that the risk associated with nursing-home expenses is more persistent than that for medical expenses and can account for 3 per cent of aggregate wealth, with half of the total effect accruing pre-retirement.

Lockwood (2016) finds that both medical and nursing-home expenses, *and* luxury-type bequests are important to understand the prevalence both of low dissaving rates and low rates of long-term care insurance coverage among wealthier retirees. Intuitively, a luxury-type bequests motive not only significantly increases saving by richer individuals as in De Nardi (2004a), it also significantly reduces the demand for insurance against late-life risk by lowering the opportunity cost of precautionary saving.

Thus, modeling these forces is an important avenue for future research to better understand their effects on wealth inequality at all ages.

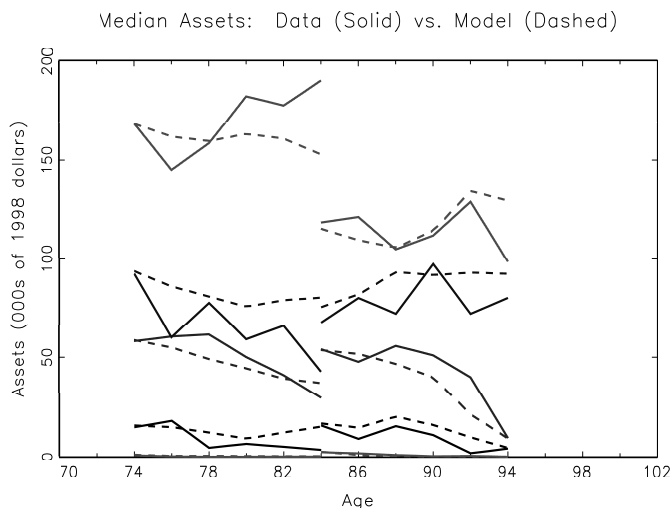


Figure 5: Savings after age 74, by age and permanent income. HRS data (solid line) and model fit (dashed line).

7.1 Interesting potential extensions

It would be interesting to evaluate the role of uncertain medical expenses at all stages of the life cycle and their implications in terms of wealth inequality. De Nardi et al. (2016b) study the effects of the costs of bad health over the whole life cycle, but do not focus on the implications of medical costs on wealth inequality at all ages and in the cross-section.

8 Heterogeneity in rates of return

The standard Bewley model assumes that all individuals are confronted with a common risk-free rate of return on saving. Yet, there is evidence that rates of return are heterogenous and often quite risky.

Earlier attempts to estimate the degree of heterogeneity in rates of return are by Flavin and Yamashita (2002), who compare the risk and return on housing to those of various asset categories and portfolios. Moskowitz and Vissing-Jørgensen (2002) find that the returns to private equity are no higher than the returns to public equity in the 1990s, while Kartashova (2014) uncovers that the difference between private and public equity returns is positive and large period-by-period between 1999 and 2007.

More recently, Bach et al. (2015) evaluate the portfolios of wealthy households in Sweden and find that yearly returns to financial wealth are on average 4% higher for households in the top 1% of the wealth distribution, compared with the median household, but that these high average returns are primarily compensations for higher levels of systematic risk. Work by Fagereng et al. (2016) uses high-quality Norwegian tax data and also provides evidence of substantial heterogeneity in individual returns to wealth.

In particular, they document: (a) a spread of 500 basis between the 10th and the 90th percentile of the distribution of returns; and (b) that heterogeneity holds within asset classes, rather than just being the result of a different portfolio mix between safe and risky assets.

The possibility that rates of return to wealth accumulation are subject to random idiosyncratic shocks has important implications for the dynamics of wealth inequality and wealth concentration. Stochastic rates of return change the technology (and more specifically the stochastic process for factor returns) in our model economy and imply that the process for individual wealth accumulation,

$$a_{t+1} = (1 + r_t^i) a_t + y_t - c_t,$$

has an idiosyncratic, stochastic growth component r_t^i , where i indexes individuals. It is well known since the work of Champernowne (1953) that (proportional) random growth processes imply, under appropriate regularity conditions, a long (Pareto) right tail. While the early contributions were purely statistical and assumed that consumption was an exogenous constant fraction of wealth, work by Benhabib et al. (2011), Benhabib et al. (2015), Aoki and Nirei (2016), and Gabaix et al. (2016) have extended this result to micro-founded models of consumption and savings.

Benhabib et al. (2016) conduct a quantitative exploration of the extent to which idiosyncratic rates of return, skewed earnings risk, and luxury bequest motives account for both the U.S. cross-sectional wealth distribution and its inter-generational wealth mobility. They structurally estimate their model and find that idiosyncratic rates of return contribute to top wealth concentration but are not sufficient to explain it. In fact, they show that saving and bequest behavior that increase with wealth (as generated by a luxury bequest motive) are both necessary and quantitatively more important to account for top wealth inequality.¹³ Idiosyncratic rates of return are, however, crucial to explain social mobility, in particular by speeding up downward mobility.

To understand the intuition for their findings, it is useful to rewrite the individual wealth accumulation identity as

$$a_{t+1} = \left(1 + r_t^i - \frac{c_t}{a_t}\right) a_t + y_t.$$

The extent to which the proportional growth rate term $(1 + r_t^i - c_t/a_t)$ is persistent, rather than i.i.d., across individuals (type-dependence) or increasing in wealth (scale-dependence) generates positive feedback, respectively, from luck and wealth levels. Higher saving rates for wealth-rich individuals—i.e., c_t/a_t decreasing in wealth—are one such

¹³Dynan et al. (2004) and Saez and Zucman (2014) document that the rich do save at higher rates.

mechanism. Scale dependence may also work through higher rates of return for wealthier people.¹⁴ Type-dependence, effectively highly persistent differences in idiosyncratic rates of return, provide a similar amplifying mechanism. Wälde (2016) calibrates the degree of ex-ante heterogeneity in the distribution of rates of return to match the evolution of the wealth distribution of the NLSY79 cohort between 1986 and 2008.

Gabaix et al. (2016) show that either scale or type-dependence is necessary for random growth models to account not only for top wealth concentration but also for the speed of change of wealth concentration observed in the data. In their absence, the speed of transition of this class of models is extremely slow. Interestingly, Bach et al. (2015) find evidence of scale dependence, while Fagereng et al. (2016) find evidence of both scale- and type-dependence—i.e., across current rate of return, level and type of education, and access to private equity.

8.1 Interesting potential extensions

In summary, persistent heterogeneity in rates of return is an important mechanism that generates at least some of the wealth inequality that we observe. Rates of return, however, are often the result of individual choices, knowledge, and ability. For instance, for entrepreneurs they are endogenous to the decision to start a business, to the amount saved, and to the share of wealth invested in the business. For investors, the rate of return is the result of their saving and portfolio choices, including housing. Lusardi et al. (Forthcoming) show that heterogeneity in rates of return driven by endogenous differences in financial knowledge can account for 30 to 40 percent of wealth inequality. Among the models studying portfolio choice and wealth inequality, Kacperczyk et al. (2015) quantitatively evaluate portfolio choice in the presence of endogenous information acquisition and heterogeneity in investor sophistication and asset riskiness. They show that an improvement in the aggregate technology to process information can explain the observed increase in wealth concentration among investors since 1990.¹⁵

This raises the important question of how rates of return, particularly at the top of the wealth distribution, are determined. More work is needed to shed additional light on the key determinants of the rates of return to one's wealth and their role in generating wealth inequality. We now turn to the study of entrepreneurship as an important determinant of rates of return and inequality in wealth holdings.

¹⁴In Kaplan et al. (2016), returns on wealth are increasing in wealth due to a fixed cost of portfolio adjustment for high-return illiquid assets. In conjunction with a perpetual youth demographic, which generates a number of individuals with very long life spans, the model generates a long top wealth tail.

¹⁵Somewhat related, Mengus and Pancrazi (2016) extend Aiyagari (1994) by allowing individuals to choose to access a complete asset market at a fixed cost instead of investing only in a risk-free asset. They show that for parameter configurations implying a non-degenerate distribution of agents over both markets, wealth inequality can be substantially higher than if the equilibrium features only investment in the risk-free asset.

9 Entrepreneurship

Quadrini (1999), Gentry and Hubbard (2004), Buera (2006), and De Nardi et al. (2007) argue that entrepreneurship is a key element generating wealth concentration among the richest households.¹⁶ Below, we present some data indicating that this is the case, but to identify what an entrepreneur is in the data, let us first see what an entrepreneur is in the model that we adopt.

Quadrini (1999) was the first to study the interaction between entrepreneurship and saving behavior by introducing an endogenous entrepreneurial choice in a Bewley model. Cagetti and De Nardi (2006) build on his contribution. They build a model of entrepreneurship with perfectly altruistic, finitely-lived agents who are endowed with two types of abilities—as a worker and as an entrepreneur—and with an entrepreneurial production function that endogenizes the rate of return to being an entrepreneur.

Cagetti and De Nardi (2006)'s model has the following key elements. Completely altruistic agents care about their children and face uncertainty about their time of death. Thus, they leave both accidental and voluntary bequests. There are two stages of life, the model period is one year, agents age stochastically from the first to the second stage and then die stochastically. An agent that dies is replaced by an offspring that inherits assets and business, if there is one, and whose abilities are correlated to those of the deceased parent. Every period, the young agents observe their ability for the period both as a worker and as an entrepreneur, form expectations about their future realizations, and decide whether to run a business or work for a wage.

The entrepreneurial production function is given by

$$f(k) = \theta k^\nu + (1 - \delta)k,$$

where k is working capital, θ is entrepreneurial ability, ν is the degree of decreasing returns to scale, and δ is depreciation. Cagetti and De Nardi (2009) generalize the entrepreneurial production function to labor hiring.

Borrowing constraints imply that

$$k = a + b(a),$$

where a is one's assets and $b(a)$ is borrowing as a function of one's assets.¹⁷

¹⁶Quadrini (2009) surveys the factors affecting the decision to become an entrepreneur and the aggregate and distributional implications of entrepreneurship for savings and investment.

¹⁷The notation does not allow for dependence on all the state variables. In the formulation adopted in Cagetti and De Nardi (2006), $b(a)$ is actually a function of all of the state variables in the economy and this outcome arises endogenously from the assumption that contracts are imperfectly enforceable and that lenders take the imperfect enforceability of contracts into account when deciding how much to lend (as in Cooley et al. (2004) and Kehoe and Levine (1993)). Besides being more micro-founded, these kind

Top %	1	5	10	20
Whole population				
Percentage of total net worth held	30	54	67	81
Entrepreneurs				
Percentage of households in a given percentile	63	49	39	28
Percentage of net worth held in a given percentile	68	58	53	47

Table 6: From Cagetti and De Nardi (2006). Entrepreneurs and the distribution of wealth. SCF 1989.

An important issue is how one identifies entrepreneurs in the data given the definition of entrepreneur adopted in the model. Cagetti and De Nardi (2006) use the SCF and classify as *entrepreneurs* those households who declare being self-employed, owning a privately held business (or a share of one), and having an active management role in it. According to this definition, which is consistent with the one in the model that they use, entrepreneurs constitute a small fraction of the population (about 8%) but hold a large share of total net worth (about 40%). They show that, in the data, entrepreneurs constitute a large fraction of rich people.

Table 6, from their paper, shows that, not only the total net worth held by the top percentiles (first row), but also the percentage of entrepreneurs in a given wealth percentile (second row) and the percentage of wealth within that percentile that is owned by entrepreneurs (third row) are all very high. For example, among the richest 1% of people in terms of net worth, 63% are entrepreneurs, and they hold 68% of the total wealth held by the wealthiest 1% of people (who hold 30% of total net worth). They also show that alternative classifications of entrepreneurship give similar results.

In Cagetti and De Nardi (2006)'s calibration, the optimal firm size is large and the entrepreneur is borrowing constrained. Thus, entrepreneurs, even when rich, want to keep saving to grow their firm to be able to borrow more and reap higher returns from capital. This is the mechanism that, in this framework, keeps the rich entrepreneurs' saving rate high and generates high wealth concentration.

In order to compare buffer-stock saving behavior with entrepreneurial saving behavior, Figure 6 compares the saving rates (defined as assets in a given period minus assets in the previous period, divided by total income during the period) for people who have the highest ability level as workers during the current period. The solid line refers to the

of borrowing constraints also have the advantage of endogenously responding to economic conditions such as changing wages and interest rates (see Bassetto et al., 2014, for an illustration and a discussion of this mechanism applied to the Great Recession). However, simpler kinds of borrowing constraints, such as linear functions of one's assets, make for models that are easier and faster to solve, and generate similar implications for cross-sectional wealth inequality at one point in time. For an application of the classic case in which borrowing is a linear function of one's assets in a model with wealth inequality and entrepreneurship, see Kitao (2008) and Meh (2005).

Wealth Gini	Fraction of entrepreneurs	Percentage wealth in the top			
		1%	5%	20%	40%
Baseline model with entrepreneurs					
0.8	7.50%	31	60	83	94

Table 7: Cagetti and De Nardi (2006) model's implications.

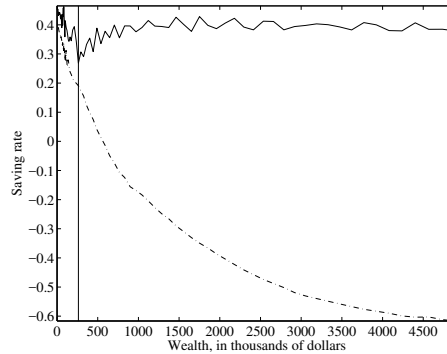


Figure 6: Saving rate for highest-ability workers. Solid line: with high entrepreneurial ability; dash-dot line: with no entrepreneurial ability; vertical line: asset level at which high-entrepreneurial-ability individuals enter entrepreneurship.

people who draw the high entrepreneurial ability level during the current period, while the dash-dotted line refers to those who get the low entrepreneurial ability draw. Given the same asset level (and potential earnings as workers), agents with high entrepreneurial ability have a much higher saving rate while workers, with no entrepreneurial ability, display pure buffer-stock saving behavior.

Agents with high entrepreneurial ability become entrepreneurs only if their wealth is above a certain threshold, corresponding to the vertical line in the figure. The saving rate of those with high entrepreneurial ability but who do not own enough assets to become entrepreneurs is higher than the one for agents without entrepreneurial ability. Intuitively, as ability is persistent, workers with high entrepreneurial ability save to have a chance to start a business in the future. In the region to the left of the threshold, the distance between the solid and the dash-dotted lines is solely due to the higher implicit rate of return from saving that one could obtain by becoming an entrepreneur in the future. All households with wealth in that range choose to be workers and earn the same income. Yet, the desire to become entrepreneurs generates a higher saving rate for agents with high entrepreneurial ability.

The saving rate of those with high entrepreneurial ability and enough assets to become entrepreneurs (in the region to the right of the threshold) is positive and considerably higher than that of workers. The return on entrepreneurial activity is high, and the entrepreneur would like to increase the size of the firm by borrowing capital. However, the

borrowing constraint limits the size of the firm and entrepreneurs must partly self-finance any additional investment. Therefore, the combination of higher returns from the business together with the budget constraint generates a high saving rate for entrepreneurs. As the firm size expands, returns eventually decrease and so does the saving rate. (We truncate the axis of the graph for easier readability.)

Table 7 shows that the model is successful in generating a high degree of wealth concentration. A few things are worth mentioning. First, the distribution of wealth is not matched by construction in the calibration procedure. Second, the model's implied returns to capital are not implausibly high and are within the range of those found by Moskowitz and Vissing-Jørgensen (2002) and Kartashova (2014). Third, the model generates entry probabilities as a function of one's wealth that are consistent with those estimated by Hurst and Lusardi (2004) on micro-level data and also implies that inheritances are a strong predictor of business entry.

In a related contribution, Herranz et al. (2015) study the interaction of heterogeneity in risk aversion and entrepreneurial firm size, capital structure, and default to manage risk, but do not study their model's implications for wealth inequality.

9.1 Interesting potential extensions

Cagetti and De Nardi (2006)'s model imposes several important restrictions: the life-cycle structure is stylized, there is only one type of entrepreneurial production function, agents can either run their business or work for a firm, and they cannot get around the borrowing constraint by selling the firm or going public.

In contrast, the data point to large heterogeneity among entrepreneurs, including in the growth and development of their firm, in their aspirations, and in their attachment to working for a firm while they also operate their own business. Campbell and De Nardi (2009) find, for instance, that aspirations about the size of the firm that one would like to run are different for men and women, and that many people who are trying to start a business also work for an employer and thus work long hours in total. Hence, many people get their business started while still working as an employee.

It would be interesting to generalize Cagetti and De Nardi (2006)'s model to allow for heterogeneity in entrepreneurial production functions. Given the data on time allocation, it would also be worthwhile to think more about the time allocation decision between working for an employer, starting and running one's firm, and home production. The challenge would be to convincingly take the additional richness in the model to the data, but this heterogeneity is clearly important for a number of questions, including the effects of taxation and government support programs on various types of entrepreneurs.

Extending Cagetti and De Nardi (2006)'s framework to a richer life-cycle structure

would allow us to better understand the timing and trade-offs of being an entrepreneur or a worker over the life cycle and to compare them for men and women who, in the data, also differ in their likelihood of entering entrepreneurship by age (Campbell and De Nardi (2009)). Women might, for instance, enter entrepreneurship later, during their child-rearing years, seeking a more convenient and flexible occupation, rather than an enterprise with a potentially high rate of return.

Finally, Cagetti and De Nardi (2006) does not allow for firms to be sold or for an explicit decision to go public. Understanding the life cycle of firms, the degree to which they are borrowing constrained as they mature, and the decision to go public or sell one's firm is important and also deserves more investigation. Glover and Short (2015) study the interplay between entrepreneurial risks and the decisions to incorporate and to go bankrupt. Chari et al. (2004) study the role of capital gains taxation on business start-up and on the sale of one's firm to professional managers. None of these papers, however, focus on their model's implications for wealth inequality.

10 Lessons learned and directions for future research

Basic versions of the Bewley model miss key aspects of saving behavior and, in particular, the saving behavior of the rich. They also overstate the importance of precautionary savings due to earnings risk, because they do not allow for the presence of any other risks or other reasons to save.

Reviewing the previous contributions, this survey finds that explicitly modeling the life cycle¹⁸ and the transmission of human capital and bequests across generations, entrepreneurship, and medical expenses is key to understanding saving and wealth inequality. These forces, however, have mostly been studied in isolation, which makes it difficult to establish their relative importance. For this reason, we argue that a rich model including these forces should be developed. As the model's richness grows, more aspects of the data, in addition to wealth inequality, need to be compared with the model's implications, to disentangle the strength and importance of the various saving motives. These additional moments should include wealth mobility, both within and across generations, for both entrepreneurs and workers, portfolio composition, the marginal propensity to consume and bequeath out of an additional dollar, and the correlation between lifetime earnings and wealth at retirement.

This survey concludes by mentioning other important economic forces that so far have not been examined in the context of quantitative models of wealth inequality and that are potentially important for future research. They include modeling the family, endogenizing wages and human capital, endogenizing health, better measuring and

¹⁸See also Krueger et al. (2016a).

understanding idiosyncratic risk, and endogenizing rates of return from assets.

First, most people are in families. Importantly, this exposes them to their own risks and their partner's risks (earnings, medical expenses, death, and divorce), but provides them with the ability to self-insure using the labor supply of both partners and savings. Blundell et al. (2012) highlight the importance of family labor supply as an insurance mechanism for wage shocks and find strong evidence of smoothing of permanent shocks to wages. Doepke and Tertilt (2016) discuss the importance of families in macroeconomics more broadly. Attanasio et al. (2015) and Borella et al. (2016) focus on risk-sharing within the family, but not on the implications for wealth inequality. Given that a significant portion of wealth is saved to self-insure against risks and that the family is an important source of both risks and insurance, it is vital to model these mechanisms and to understand their effects on wealth inequality.

Second, it is important to jointly model the determinants and dynamics of both earnings and wealth over the life cycle by allowing for endogenous human capital accumulation, in addition to savings. Huggett et al. (2006) show that a benchmark model of human capital can replicate mean earnings and measures of earnings dispersion and skewness over the working life cycle observed in the data. Huggett et al. (2011) find that, in a model of risky human capital, as of age 23, differences in initial conditions account for more of the variation in lifetime earnings, lifetime wealth, and lifetime utility than do differences in shocks received over the working lifetime. These findings raise the question of what are the implications of human capital and how its accumulation interacts with that of wealth, both over the life cycle and in the cross-section.

Third, it is essential to understand the role of health dynamics and medical expenses in determining labor supply, earnings, and savings. De Nardi et al. (2010) show that medical expenses and lifespan heterogeneity that are heterogenous by income quintiles are important to explain saving behavior by wealth quintile during old age. It would be interesting to embed medical expenses in a full life-cycle model and analyze their implications for wealth inequality. Going one step backwards, one should think about the determinants of health and medical expenditures and ideally model jointly inequality in health, wealth, and medical expenses over all of the life cycle.

Fourth, the importance of the nature of idiosyncratic risk assumed in these models also raises the question of its measurement in the data. What we, as economists, measure as a shock in the data might be anticipated by the households. This might be especially true for administrative data sets that contain little information about the household (in contrast with survey data, which instead, might contain information on households' health, divorce, and expectations). Sabelhaus and Ackerman (2012) use SCF data to derive the gap between actual and normal income from survey questions and use it as a measure of shocks. This approach stands in contrast to existing income shock measures

in the literature, which are generally derived from the residuals of estimated earnings or income equations. Interestingly, the overall variance and asymmetry of shocks over the business cycle derived from this analysis are similar to those of existing residual-based estimates. Blundell et al. (2008) use data on both consumption and income to draw inference on the persistence of income shocks. More work is needed to better disentangle the actual shocks that households face and their sources.

Fifth, while heterogeneity in rates of return plays some role in generating wealth inequality, as shown by Benhabib et al. (2016), and while there is evidence that rates of return are heterogenous, it is important to recognize that they are endogenous and dependent on occupational choice or entrepreneurship and portfolio choice. Much more work is needed, to determine what gives rise to the heterogeneity in rates of return that we observe and how it interacts with wealth accumulation and wealth inequality.

It is important to note that understanding these mechanisms is also crucial from the standpoint of studying the implications of different government policies. More specifically, different mechanisms may give rise to similar observed wealth concentrations but have vastly different policy implications. For instance, modeling entrepreneurship often implies that the adverse responses of savings and economic activity to increased taxation are significant, and especially so if taxation affects the returns to running a business (Kitao (2008), Lee (2015), and Cagetti and De Nardi (2009)). In contrast, in a model with high earnings risk for the top earners, Kindermann and Krueger (2015) conclude that the optimal marginal income tax rate is close to 90%. The big difference in responses to taxation between these models is due to the fact that entrepreneurs' savings and investments are responsive to their implicit rate of return, net of taxes. In contrast, individuals with very high labor earnings facing a large probability of a very large fall in earnings next period have very high incentives to engage in precautionary saving. Hence, when an increase in labor taxation reduces their net earnings, they still save at a high rate, as long as expected net income tomorrow is sufficiently low compared with today's net earnings. This stark contrast in policy implications stemming from different motivations to save points to the importance of understanding whether, for instance, the risk that the rich face comes from the return on their human capital as opposed to the return on the wealth they have invested in their business. More work needs to be done to more conclusively determine the effects of taxation in quantitative models of wealth inequality.

Finally, in this survey, we focus on the determinants of inequality at a point in time, but more work is needed to understand the dynamics of inequality and its determinants over time.¹⁹ For some work on the evolution of wealth inequality over time, see Hubmer

¹⁹For an interesting model of how consumption and income (but not wealth) inequality evolve over time, see Krueger and Perri (2006). For empirical papers studying changes of earnings processes over time, see, among others, DeBacker et al. (2013), Sabelhaus and Song (2009), and Dynan et al. (2007).

et al. (2016), Kaymak and Poschke (2016), and Gabaix et al. (2016).

References

- John M. Abowd and David Card. On the covariance structure of earnings and hours changes. *Econometrica*, 57:441–455, March 1989.
- S. Rao Aiyagari. Uninsured idiosyncratic risk and aggregate saving. *Quarterly Journal of Economics*, 109(3):659–684, August 1994.
- S Rao Aiyagari, Jeremy Greenwood, and Ananth Seshadri. Efficient investment in children. *Journal of Economic Theory*, 102(2):290–321, 2002.
- Joseph G. Altonji and Ernesto Villanueva. The marginal propensity to spend on adult children. NBER Working Paper 9811, July 2003.
- James Andreoni. Giving with impure altruism: Applications to charity and Ricardian equivalence. *Journal of Political Economy*, 97:1447–1458, 1989.
- Shuhei Aoki and Makoto Nirei. Pareto distribution of income in neoclassical growth models. *Review of Economics Dynamics*, 20:25–42, 2016.
- Manuel Arellano, Richard Blundell, and Stéphane Bonhomme. Earnings and consumption dynamics: A non-linear panel data framework. *Econometrica*, 2017. forthcoming.
- Orazio Attanasio, Hamish Low, and Virginia Sanchez-Marcos. Female labor supply as insurance against idiosyncratic risk. *Journal of the European Economic Association*, 3: 755–764, 2015.
- Laurent Bach, Laurent E. Calvet, and Paolo Sodini. Rich pickings? risk, return, and skill in the portfolios of the wealthy. Mimeo HEC Paris, 2015.
- Michael Barnett and Vasia Panousi. Entrepreneurial risk and the great recession. Mimeo, 2015.
- Marco Bassetto, Marco Cagetti, and Mariacristina De Nardi. Credit crunches and credit allocation in a model of entrepreneurship. *Review of Economic Dynamics*, 2014.
- Gary Becker and Nigel Tomes. Human capital and the rise and fall of families. *Journal of Labor Economics*, 4(3):S1–S39, July 1979.
- Jess Benhabib, Alberto Bisin, and Shenghao Zhu. The distribution of wealth and fiscal policy in economies with finitely lived agents. *Econometrica*, 79(1):123–157, 2011.
- Jess Benhabib, Alberto Bisin, and Shenghao Zhu. The wealth distribution in Bewley models with capital income risk. *Journal of Economic Theory*, 159:459–515, 2015.

- Jess Benhabib, Alberto Bisin, and Mi Luo. Wealth distribution and social mobility in the US: A quantitative approach. Mimeo, 2016.
- B. Douglas Bernheim, Jonathan Skinner, and Steven Weinberg. What accounts for the variation in retirement wealth among U.S. households? *American Economic Review*, 91(4):832–857, September 2001.
- Truman F. Bewley. The permanent income hypothesis: A theoretical formulation. *Journal of Economic Theory*, 16(2):252–292, December 1977.
- Richard Blundell, Luigi Pistaferri, and Ian Preston. Consumption inequality and partial insurance. *The American Economic Review*, 98(5):1887–1921, 2008.
- Richard Blundell, Luigi Pistaferri, and Itay Saporta-Eksten. Consumption inequality and family labor supply. NBER Working Paper 18445, October 2012.
- Richard Blundell, Michael Graber, and Magne Mogstad. Labor income dynamics and the insurance from taxes, transfers and the family. *Journal of Public Economics*, 127: 58–73, 2015.
- Stéphane Bonhomme and Jean-Marc Robin. Assessing the equalizing force of mobility using short panels: France, 1990–2000. *The Review of Economic Studies*, 76(1):63–92, 2009.
- Margherita Borella, Mariacristina De Nardi, and Fang Yang. Risks and insurance over the life cycle for couples and singles. UCL, 2016.
- Samuel Bowles, Herbert Gintis, and Melissa Osborne Groves. *Unequal chances: Family background and economic success*. Princeton University Press, 2009.
- Jesse Bricker, John Sabelhaus, Jake Krimmel, and Alice Henriques. Measuring income and wealth at the top using administrative and survey data. Mimeo, 2015.
- Meta Brown, John Karl Scholz, and Ananth Seshadri. A new test of borrowing constraints for education. *The Review of Economic Studies*, page rdr032, 2011.
- Martin Browning and Annamaria Lusardi. Household saving: Micro theories and micro facts. *Journal of Economic Literature*, 24:1797–1855, December 1996.
- Francisco Buera. Persistency of poverty, financial frictions, and entrepreneurship. Working paper, Northwestern University, 2006.
- Francisco Buera. A dynamic model of entrepreneurship with borrowing constraints. *Annals of Finance*, 5(3-4):443–464, 2009.

- Marco Cagetti. Wealth accumulation over the life cycle and precautionary savings. *Journal of Business and Economic Statistics*, 21(3):339–353, July 2003.
- Marco Cagetti and Mariacristina De Nardi. Entrepreneurship, frictions, and wealth. *Journal of Political Economy*, 114(5):835–870, October 2006.
- Marco Cagetti and Mariacristina De Nardi. Wealth inequality: data and models. *Macroeconomic Dynamics*, 12:285–313, 2008.
- Marco Cagetti and Mariacristina De Nardi. Estate taxation, entrepreneurship, and wealth. *The American Economic Review*, 99(1):85–111, 2009.
- Jeffrey R. Campbell and Mariacristina De Nardi. A conversation with 590 entrepreneurs. *Annals of Finance*, 5:313–327, 2009.
- Christopher D. Carroll. Buffer stock saving and the life-cycle/permanent income hypothesis. *Quarterly Journal of Economics*, 112(1):1–55, February 1997.
- Christopher D. Carroll. Why do the rich save so much? In Joel B. Slemrod, editor, *Does Atlas Shrug? The Economic Consequences of Taxing the Rich*, pages 466–484. Harvard University Press, Cambridge, 2000.
- Ana Castañeda, Javier Díaz-Giménez, and José-Victor Ríos-Rull. Accounting for the U.S. earnings and wealth inequality. *Journal of Political Economy*, 111(4):818–857, August 2003.
- D. G. Champernowne. A model of income distribution. *Economics Journal*, 63:318–351, 1953.
- V. V. Chari, M. Golosov, and Aleh Tsyvinski. Business start-ups, the lock-in effect, and capital gains taxation. Working paper, Princeton University, 2004.
- Juan Carlos Conesa, Sagiri Kitao, and Dirk Krueger. Taxing capital: Not a bad idea after all! *American Economic Review*, 99(1):25–48, 2009.
- Thomas Cooley, Ramon Marimon, and Vincenzo Quadrini. Aggregate consequences of limited contract enforceability. *Journal of Political Economy*, 112(4):817–847, August 2004.
- James B. Davies and Anthony F. Shorrocks. The distribution of wealth. In A.B. Atkinson and F. Bourguignon, editors, *Handbook of Income Distribution*, pages 605–675. Elsevier, 2000.

- Jason DeBacker, Vasia Panousi, and Shanthi Ramnath. The properties of income risk in privately held businesses. Federal Reserve Board Working Paper 2012-69, 2012.
- Jason DeBacker, Bradley Heim, Vasia Panousi, Shanthi Ramnath, and Ivan Vidangos. Rising inequality: Transitory or persistent? new evidence from a panel of u.s. tax returns. Mimeo, 2013.
- Antonia Díaz, Josep Pijoan-Mas, and José-Victor Ríos-Rull. Habit formation: Implications for the wealth distribution. *Journal of Monetary Economics*, 50(6):1257–1291, August 2002.
- Javier Díaz-Giménez, Vincenzo Quadrini, and José-Victor Ríos-Rull. Dimensions of inequality: Facts on the U.S. distributions of earnings, income and wealth. *Federal Reserve Bank of Minneapolis Quarterly Review*, 21(2):3–21, Spring 1997.
- Matthias Doepke and Michele Tertilt. Families in macroeconomics. *Forthcoming as Chapter 23 in the Handbook of Macroeconomics, Vol. 2, edited by John Taylor and Harald Uhlig*, 2016.
- Mariacristina De Nardi. Wealth inequality and intergenerational links. *Review of Economic Studies*, 71:743–768, 2004a.
- Mariacristina De Nardi. Wealth inequality and intergenerational links. *Review of Economic Studies*, 71(3):743–768, July 2004b.
- Mariacristina De Nardi and Fang Yang. Bequests and heterogeneity in retirement wealth. *European Economic Review*, 72:182–196, 2014.
- Mariacristina De Nardi and Fang Yang. Wealth inequality, family background, and estate taxation. Working Paper 21047, National Bureau of Economic Research, 2015.
- Mariacristina De Nardi, Phil Doctor, and Spencer D. Krane. Evidence on entrepreneurs in the United States: Data from the 1989-2004 Survey of Consumer Finances. *Economic Perspectives*, 4:18–36, 2007.
- Mariacristina De Nardi, Eric French, and John B. Jones. Why do the elderly save? The role of medical expenses. *Journal of Political Economy*, 118(1):39–75, 2010.
- Mariacristina De Nardi, Giulio Fella, and Gonzalo Paz-Pardo. The implications of richer earnings dynamics for consumption and wealth. NBER Working paper No, 21917, 2016a.
- Mariacristina De Nardi, Svetlana Pashchenko, and Pompoje Porapakkarm. The lifetime costs of bad health. Mimeo, University of Georgia, 2016b.

- Karen Dynan, Jonathan Skinner, and Stephen Zeldes. Do the rich save more? *Journal of Political Economy*, 112(2):397–444, April 2004.
- Karen E Dynan, Douglas W Elmendorf, Daniel E Sichel, et al. *The evolution of household income volatility*. Division of Research & Statistics and Monetary Affairs, Federal Reserve Board, 2007.
- Martin Eichenbaum and Dan Peled. Capital accumulation and annuities in an adverse selection economy. *Journal of Political Economy*, 95(2):334–354, April 1987.
- Andreas Fagereng, Luigi Guiso, Davide Malacrino, and Luigi Pistaferri. Heterogeneity and persistence in returns to wealth. Mimeo, 2016.
- Marjorie Flavin and Takashi Yamashita. Owner-occupied housing and the composition of the household portfolio. *American Economic Review*, 92(1):345–362, 2002.
- Xavier Gabaix, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll. The dynamics of inequality. *Econometrica*, forthcoming, 2016.
- William G. Gale and John Karl Scholz. Intergenerational transfers and the accumulation of wealth. *Journal of Economic Perspectives*, 8(4):145–160, Fall 1994.
- William M. Gentry and R. Glenn Hubbard. Entrepreneurship and household savings. *Advances in Economic Analysis and Policy*, 4(1), 2004. Article 1.
- John Geweke and Michael Keane. An empirical analysis of earnings dynamics among men in the psid: 1968-1989. *Journal of Econometrics*, 96:293–356, 2000.
- Andy Glover and Jacob Short. Bankruptcy, incorporation, and the nature of entrepreneurial risk. Mimeo, 2015.
- Jagadeesh Gokhale, Laurence J. Kotlikoff, James Sefton, and Martin Weale. Simulating the transmission of wealth inequality via bequests. *Journal of Public Economics*, 79(1):93–128, 2001.
- Fatih Guvenen. *Quantitative Economics With Heterogeneity: An A-to-Z Guidebook*. Princeton University Press, Princeton, New Jersey, 2016.
- Fatih Guvenen, Fatih Karahan, Serdar Ozcan, and Jae Song. What do data on millions of U.S. workers reveal about life-cycle earnings risk? NBER Working Paper 20913, January 2015a.
- Fatih Guvenen, Fatih Karahan, Serdar Ozkan, and Jae Song. What do data on millions of U.S. workers reveal about life-cycle earnings risk? Working paper, University of Minnesota, 2015b.

- Gary D Hansen and Ayşe İmrohoroğlu. The role of unemployment insurance in an economy with liquidity constraints and moral hazard. *Journal of Political Economy*, pages 118–142, 1992.
- Jonathan H. Heatcote, Kjetil Storesletten, and Giovanni L. Violante. Quantitative macroeconomics with heterogeneous households. *Annual Review of Economics*, 1:319–354, September 2009.
- James J. Heckman and Stefano Mosso. The economics of human development and social mobility. NBER Working Paper 19925, February 2014.
- Burkhard Heer. Wealth distribution and optimal inheritance taxation in life-cycle economies with intergenerational transfers. Mimeo. University of Cologne, Germany, 1999.
- Lutz Hendricks. Accounting for patterns of wealth inequality. Mimeo. Iowa State University, 2004.
- Lutz Hendricks. Retirement wealth and lifetime earnings. *International Economic Review*, 48(2):421–456, 2007a.
- Lutz Hendricks. How important is preference heterogeneity for wealth inequality? *Journal of Economic Dynamics and Control*, 31(9):3042–3068, 2007b.
- Neus Herranz, Stefan Krasa, and Anne P. Villamil. Entrepreneurs, risk aversion, and dynamic firms. *Journal of Political Economy*, 123(5):1133–1176, 2015.
- R. Glenn Hubbard, Jonathan Skinner, and Stephen P. Zeldes. The importance of precautionary motives in explaining individual and aggregate saving. *Carnegie-Rochester Conference Series on Public Policy*, 40:59–152, 1994.
- R. Glenn Hubbard, Jonathan Skinner, and Stephen P. Zeldes. Precautionary saving and social insurance. *Journal of Political Economy*, 103(2):360–399, 1995.
- Joachim Hubmer, Per Krusell, and Antony A. Smith Jr. The historical evolution of the wealth distribution: A quantitative-theoretic investigation. NBER working paper no. 23011, 2016.
- Mark Huggett. Wealth distribution in life-cycle economies. *Journal of Monetary Economics*, 38(3):469–494, December 1996.
- Mark Huggett, Gustavo Ventura, and Amir Yaron. Human capital and earnings distribution dynamics. *Journal of Monetary Economics*, 53:265–290, 2006.

- Mark Huggett, Gustavo Ventura, and Amir Yaron. Sources of lifetime inequality. *American Economic Review*, 101:2923–54, 2011.
- Michael D. Hurd. Mortality risk and bequests. *Econometrica*, 57(4):779–813, July 1989.
- Erik Hurst and Annamaria Lusardi. Liquidity constraints, wealth accumulation and entrepreneurship. *Journal of Political Economy*, 112(2):319–347, April 2004.
- Erik Hurst, Ming Ching Luoh, and Frank P. Stafford. Wealth dynamics of American families, 1984-94. *Brookings Papers on Economic Activity*, 1998(1):267–337, 1998.
- Marcin Kacperczyk, Jaromir Nosal, and Luminita Stevens. Investor sophistication and capital income inequality. Mimeo Boston College, 2015.
- Greg Kaplan, Benjamin Moll, and Giovanni L. Violante. Monetary policy according to HANK. mimeo, 2016.
- Katya Kartashova. The returns to entrepreneurial investment: A private equity premium puzzle? *American Economic Review*, 104(10):3297–3394, 2014.
- Barış Kaymak and Markus Poschke. The evolution of wealth inequality over half a century: the role of taxes, transfers and technology. *Journal of Monetary Economics*, 77:1–25, 2016.
- Timothy J. Kehoe and David K. Levine. Debt-constrained asset markets. *Review of Economic Studies*, 60(4):865–888, October 1993.
- Arthur B. Kennickell. A rolling tide: Changes in the distribution of wealth in the U.S., 1989-2001. Mimeo, September 2003.
- Fabian Kindermann and Dirk Krueger. High marginal tax rates on the top 1%? Lessons from a life cycle model with idiosyncratic income risk. Mimeo, January 2015.
- Sagiri Kitao. Entrepreneurship, taxation, and capital investment. *Review of Economic Dynamics*, 11(1):44–69, March 2008.
- Wojciech Kopczuk. What do we know about evolution of top wealth shares in the United States? NBER Working Paper 20734, December 2014.
- Wojciech Kopczuk and Joseph Lupton. To leave or not to leave: The distribution of bequest motive. *The Review of Economic Studies*, 74(1):207–235, 2007.
- Karen Kopecky and Tatyana Koreshkova. The impact of medical and nursing home expenses on savings. *American Economic Journal: Macroeconomics*, 6(3):29–72, 2014.

- Laurence J. Kotlikoff and Lawrence H. Summers. The role of intergenerational transfers in aggregate capital accumulation. *Journal of Political Economy*, 89(4):706–732, August 1981.
- Dirk Krueger and Jesus Fernandez-Villaverde. Consumption and saving over the life cycle: How important are consumer durables? *Macroeconomic Dynamics*, 15:725–770, 2011.
- Dirk Krueger and Fabrizio Perri. Does income inequality lead to consumption inequality? Evidence and theory. *The Review of Economic Studies*, 73(1):163–193, 2006.
- Dirk Krueger, Kurt Mitman, and Fabrizio Perri. Macroeconomics and household heterogeneity. *Forthcoming in The Handbook of Macroeconomics*, 2016a.
- Dirk Krueger, Kurt Mitman, and Fabrizio Perri. On the distribution of the welfare losses of large recessions. *Handbook of Macroeconomics*, forthcoming, 2016b.
- Per Krusell and Anthony Smith, Jr. Income and wealth heterogeneity in the macroeconomy. *Journal of Political Economy*, 106(5):867–896, October 1998.
- Per Krusell and Anthony Smith, Jr. Quantitative macroeconomic models with heterogeneous agents. In *Advances in Economics and Econometrics: Theory and Applications, Ninth World Congress*, pages 110–115. Elsevier, 2006.
- Moritz Kuhn and José-Víctor Ríos-Rull. 2013 update on the U.S. earnings, income and wealth distributional facts: a view from macroeconomic modelers. *Federal Reserve Bank of Minneapolis Quarterly Review*, 37(1), 2016.
- John Laitner. Secular changes in wealth inequality and inheritance. *The Economic Journal*, 111(474):691–721, 2001.
- John Laitner and Thomas F. Juster. New evidence on altruism, a study of TIAA-CREF retirees. *American Economic Review*, 86(4):893–908, 1996.
- Emily Lawrence. Poverty and the rate of time preference: Evidence from panel data. *Journal of Political Economy*, 99:54–77, 1991.
- Sang Yoon Lee. Entrepreneurs, managers and inequality. Mimeo, 2015.
- Sang Yoon Lee, Nicolas Roys, and Ananth Seshadri. The causal effect of parental human capital on childrens human capital. Technical report, mimeo, University of Wisconsin-Madison, 2014.

- Sang Yoon Tim Lee and Ananth Seshadri. On the intergenerational transmission of economic status. *Unpublished manuscript, University of Wisconsin–Madison, Department of Economics*, 2014.
- Lars Ljungqvist and Thomas J. Sargent. *Recursive Macroeconomic Theory*. MIT Press, Boston, MA, 2000.
- Lee Lockwood. Incidental bequests: Bequest motives and the choice to self-insure late-life risks. Mimeo, 2016.
- Annamaria Lusardi, Pierre-Carl Michaud, and Olivia S. Mitchell. Optimal financial knowledge and wealth inequality. *Journal of Political Economy*, Forthcoming.
- Costas Meghir and Luigi Pistaferri. Income variance dynamics and heterogeneity. *Econometrica*, 72(1):1–32, 2004.
- Cesaire Meh. Entrepreneurship, wealth inequality, and taxation. *Review of Economic Dynamics*, 8(3):688–719, July 2005.
- Eric Mengus and Roberto Pancrazi. Endogenous partial insurance and inequality. Mimeo, 2016.
- Tobias J. Moskowitz and Annette Vissing-Jørgensen. The returns to entrepreneurial investment: A private equity premium puzzle? *American Economic Review*, 92(4):745–778, 2002.
- Shinichi Nishiyama. Bequests, inter vivos transfers, and wealth distribution. *Review of Economic Dynamics*, 5(4):892–931, October 2002.
- Jonathan A. Parker and Annette Vissing-Jørgensen. Who bears aggregate fluctuations and how? *American Economic Review*, 99(2):399–405, 2009.
- Gonzalo Paz-Pardo. Preference heterogeneity and wealth inequality. mimeo UCL, 2016.
- Thomas Piketty. *Capital in the Twenty-First Century*. Harvard University Press, Cambridge, 2014.
- Vincenzo Quadrini. The importance of entrepreneurship for wealth concentration and mobility. *Review of Income and Wealth*, 45(1):1–19, March 1999.
- Vincenzo Quadrini. Entrepreneurship in macroeconomics. *Annals of Finance*, 5(3):295–311, 2009.
- Vincenzo Quadrini and José-Victor Ríos-Rull. Understanding the u.s. wealth distribution. *Federal Reserve Bank of Minneapolis Quarterly Review*, 21(2):1–21, Spring 1997.

- Vincenzo Quadrini and José-Victor Ríos-Rull. Inequality in macroeconomics. In A.B. Atkinson and F. Bourguignon, editors, *Handbook of Income Distribution, vol. 2B*. Elsevier, 2014.
- Santiago Budria Rodriguez, Javier Díaz-Giménez, Vincenzo Quadrini, and José-Victor Ríos-Rull. Updated facts on the U.S. distributions of earnings, income, and wealth. *Federal Reserve Bank of Minneapolis Quarterly Review*, 26(3):2–35, Summer 2002.
- John Sabelhaus and Samuel Ackerman. The effect of self-reported transitory income shocks on household spending. *FEDS Working Paper*, 2012.
- John Sabelhaus and Jae Song. Earnings volatility across groups and time. *Available at SSRN 1374970*, 2009.
- Emanuel Saez and Gabriel Zucman. Wealth inequality in the United States since 1913: Evidence from capitalized income tax data. NBER Working Paper 20625, October 2014.
- John Karl Scholz and Ananth Seshadri. Children and household wealth. *Michigan Retirement Research Center Research Paper No. WP*, 158, 2007.
- Steven F. Venti and David A. Wise. The cause of wealth dispersion at retirement: Choice or chance? *American Economic Review*, 88(2):185–191, May 1998.
- Klaus Wälde. Pareto-improving redistribution of wealth – the case of the nlsy 1979 cohort. mimeo, 2016.
- Chong Wang, Neng Wang, and Jinqiang Yang. Optimal consumption and savings with stochastic income and recursive utility. Mimeo, January 2015.
- Edward N. Wolff. Changing inequality of wealth. *American Economic Review*, 82(2): 552–558, may 1992.
- Edward N. Wolff. Recent trends in the size distribution of household wealth. *Journal of Economic Perspectives*, 12(3):131–150, Summer 1998.
- Edward N Wolff and Maury Gittleman. Inheritances and the distribution of wealth or whatever happened to the great inheritance boom? *The Journal of economic inequality*, 12(4):439–468, 2014.