The distribution of wealth:

Intergenerational transmission

and redistributive policies[*]

Jess Benhabib

Department of Economics

New York University

Alberto Bisin

Department of Economics

New York University

This draft: September 2007

Abstract

We study the dynamics of the distribution of wealth in an economy with inter-generational transmission of wealth and redistributive fiscal policy. We characterize the transitional dynamics of the distribution of wealth as well as its stationary state. We show that the stationary wealth distribution is a Pareto distribution.We

study analytically the dependence of the distribution of wealth, of wealth inequality, and of utilitarian social welfare on various redistributive fiscal policy instruments like capital income taxes, estate taxes, and welfare subsidies.

1 Introduction

Rather invariably across a large cross-section of countries and time periods income and wealth distributions are skewed to the right and display a heavy upper tail (slowly declining top wealth shares). These observations have lead Vilfredo Pareto, in the *Cours d'Economie Politique* (1897), to introduce the distributions which take his name[1] and to theorize about the possible economic and sociological factors generating wealth distributions of such form. The results of Pareto's investigations take the form of the "Pareto's Law," enunciated e.g., by Samuelson (1965) as follows:

> In all places and all times, the distribution of income remains the same. Neither institutional change nor egalitarian taxation can alter this fundamental constant of social sciences.

Since Pareto, economists have lost confidence in "fundamental constant(s) of social sciences".[2] Nonetheless distributions of income and wealth which are very concentrated and skewed to the right have been well documented over time and across countries. For example, Atkinson (2001), Moriguchi-Saez (2005), Piketty (2001), Piketty-Saez (2003), and Saez-Veall (2003) document skewed distributions of income, including capital income, with relatively large top shares consistently over the last century, respectively, in the U.K., Japan, France, the U.S., and Canada. Large top wealth shares in the U.S.

---

[1] Pareto distributions are power laws. They display heavy tails, in the sense that the frequency of events in the tails of the distribution declines more slowly than e.g., in a Normal distribution. They represent a subset of the class of stable Levy distributions, that is, of the distributions which are obtained from the version of the Central Limit Theorem which does not impose finite mean and variance; see e.g., Nolan (2005).

[2] See Chipman (1976) for a discussion on the controversy between Pareto and Pigou regarding the interpretation of the Law. To be fair to Pareto, his view was not necessarily that fiscal policies cannot alter the distribution of wealth, but that fiscal policy is determined by the controlling elites who use it to skew the distribution to their advantage; see Pareto (1900).

since the 60's are documented e.g., by Wolff (1987, 2004).[3] Also, heavy upper tails of the distributions of income and wealth is a well documented empirical regularity; e.g. Klassa, Bihama, Levy, Malcaia and Solomon (2006) for the US, Nirei-Souma (2004) for the U.S. and Japan from 1960 to 1999, Clementi-Gallegati (2004) for Italy from 1977 to 2002, and Dagsvik-Vatne (1999) for Norway in 1998.

While Pareto was skeptical that "egalitarian taxation" could have any significant effect on the distribution of income, many have later concluded that the redistributive taxation regimes introduced after World War II did in fact significantly reduce income and wealth inequality; notably, e.g., Lampman (1962) and Kuznets (1955). Most recently, Piketty (2001) has argued that redistributive taxation may have prevented large income shares from recovering after the shocks that they experienced during World War II in France.[4]

In this paper we study the dynamics of the distribution of wealth as generated by inter-generational transmission and redistributive fiscal policy. We model inter-generational transmission of wealth is induced by parental altruism, in the form either of pure altruism or of "joy of giving" preferences for bequests. Redistributive fiscal policy is implemented through welfare subsidies financed by capital income taxes and estate taxes on bequeathed wealth.

More specifically, our economy is populated by a continuum of age structured overlapping generations of agents with a constant probability of death as in Blanchard (1985) and Yaari (1965). The population is stationary and each agent who dies is substituted

---

[3]While income and wealth are correlated and have qualitatively similar distributions, wealth tends to be more concentrated than income. For instance the Gini coefficient of the distribution of wealth in the U.S. in 1992 is .78, while it is only .57 for the distribution of income (Diaz Gimenez-Quadrini-Rios Rull, 1997); see also Feenberg-Poterba (2000).

[4]This line of argument has been extended to the U.S., Japan, and Canada, respectively, by Piketty-Saez (2003) and Moriguchi-Saez (2005), Saez-Veall (2003).

by his/her child. A fraction of the agents are altruistic towards children and optimally choose the amount of bequests they leave. Agents are born with an initial wealth which is composed of the present discounted value of their expected future earnings, the bequests of their parents and, if they qualify, welfare subsidies from the government. Agents face a constant interest rate. They choose an optimal consumption-savings plan, which includes the allocation of their wealth between annuities and assets which are bequeathed at their death. The government taxes capital income and estates to redistribute wealth in the form of welfare subsidies. The government budget is balanced.

We show that our model generates a a stationary wealth distribution which has the main qualitative properties which characterize wealth distributions: skewedness and fat tails. To stress inter-generational transmission as the process of wealth dynamics which we study, we disregard agents' earnings in most of the formal analysis.[5] In other words, we study *old money* rather than *new money*.[6] In particular, we show that inter-generational transmission and redistributive fiscal policy by themselves, without earnings' heterogeneity, are sufficient to induce a stationary wealth distribution which is a power law, a Pareto distribution in particular. We are therefore able to characterize the wealth distribution in closed form and hence to perform various comparative statics

---

[5]The importance of intergenerational transfers in accounting for wealth accumulation has emphasized by Kotlikoff and Summers (1981). They argue that intergenerational wealth transfers, rather than life cycle earnings, account for up to 80% of wealth accumulation. See also Gale and Scholtz (1994) for more moderate findings along the same lines.

[6]The relative importance of *old money* vs *new money* has been studied by Elwood, Miller, Bayard, Watson, Collins, and Hartman (1999). They classify the wealthiest individuals in the Forbes 400 in 1995 and 1996 according to whether they represent old or new money. They find that 43.5% of those on the list came from old money, that is they inherited sufficient wealth to rank among Forbes 400, while 30.1% represented new money, consisting of individuals and families whose parents did not have great wealth or own a business with more than a few employees. The remaining 26.4% were in intermediate categories. See also Burris (2000), p. 364, footnote 3. Finally, see Piketty-Saez (2003) for evidence of the relative decline of *old money* in accounting for top income shares in the U.S. since the 60's.

exercises which enable us to analytically study the determinants of the wealth distribution and, in particolar, of wealth inequality. The two critical ingredients that drive the Pareto wealth distribution in our model are $i)$ the accumulation of wealth with time and, most importantly, through inheritance, and $ii)$ the redistribution of wealth to the young poor through estate and capital taxes. The level of concentration and of inequality of wealth at the stationary distribution depends on the demographic characteristics of the economy, its structural parameters, as well as on the endogenous growth rate of the economy. Most specifically, wealth is less concentrated (the Gini coefficient is lower) the higher is the density of agents receiving welfare subsidies, that is, the more wealth is redistributed via welfare subsidies. Furthermore, wealth is less concentrated the lower is the growth rate of individual wealth accumulation and the higher is the growth rate of aggregate wealth. The wedge between the individual and the aggregate growth rates of wealth depends in turn also on fiscal policy: the more redistributive is fiscal policy, the smaller is the wedge, and hence the less concentrated in the wealth in the economy.

Furthermore, our explicit characterization of the stationary distribution of wealth allows us to study analytically the dependence of wealth inequality on the different redistributive fiscal policy instruments we study, capital income taxes and estate taxes. In particular, wealth is less concentrated for both higher capital income taxes and estate taxes, but the marginal effect of capital income taxes is much stronger than the effect of estate taxes. When earnings are entered in the analysis we cannot solve for the whole wealth distribution in closed form. However, we show that our results hold unchanged for the tail of the wealth distribution. In other words, it is the intergenerational transmission of wealth which in our model determines the form of the tail of the distribution. We also perform a tentative calibration exercise to illustrate the possible contribution of the inter-generational transmission of wealth to the determination of the observed levels of wealth inequality. We find that the wealth inequality induced by intergenerational transmission accounts for just little less than a third of the wealth inequality in the U.S.

6

in 1992. Finally, we characterize optimal redistributive taxes with respect to a utilitarian social welfare measure. With such an "egalitarian" welfare measure, in an economy with no labor earnings,[7] maximizing social welfare is almost equivalent to minimizing the concentration or inequality of wealth. Nonetheless, completely minimizing wealth inequality requires reducing the economy's growth rate, hence at the optimal taxes inequality is not eliminated, but in our rough benchmark calibration growth would be reduced by 1.1%. Concentrating on policies which directly affect the intergenerational transmission of wealth, we show that, robustly around our benchmark calibrated economy, social welfare is maximized with zero estate taxes. Social welfare maximizing capital income taxes, on the contrary, are positive and close to the value which minimizes wealth inequality. Our social welfare function assigns positive weight to agents currently alive, and future generations enter into the social welfare computations only through the bequest motive. If we weighted future generations separately as well, we would be putting more weight on growth, so that the optimal taxes would tolerate greater inequality and would penalize growth by less.

1.1 Related literature

A large and diverse theoretical literature on the dynamics of individual wealth dating back to the 1950es obtains distributions exhibiting power laws and, in particular, Pareto distributions. Notably, Champernowne (1953), Rutherford (1955), Simon (1955), and most of the subsequent literature study accumulation models in which an exogenous stochastic process drives wealth accumulation differentially for low and high wealth ranges. Typically in these models the stochastic processes are such that there is a lower a reflective barrier to wealth, and the higher levels of wealth thinned out by death, or by negative expected growth. Wold-Whittle (1957) in particular study a birth and death process with

---

[7]Labor earnings would naturally limit the scope of equalization measures which reduce the agents' incentives to supply labor or invest in human capital.

population growth, exogenous exponential wealth accumulation, and bequests. Most recently, the analysis of stochastic processes generating power laws in the distribution of wealth has become an important subject in Econophysics (see Mantegna-Stanley, 2000, Gabaix-Gopikrishnan-Plerou-Stanley 2003).[8]

The stochastic processes which generate power laws in this whole literature are essentially exogenous, that is, they are not the result of agents' optimal consumption-savings decisions. The dynamics of wealth in these models, therefore, is not related to the deep structural parameters of the economy nor to any policy parameter of interest. It is then impossible in the context of these models to pursue one of the objectives of this paper, that is, to study the dependence of the distribution of wealth on fiscal policy. In fact a specific fiscal policy affects the distribution of wealth in equilibrium not only through its direct redistributive effects, but also through its indirect effect on the economy's aggregate growth rate, the rate of accumulation of private savings, and through its effect on other fiscal policies required by government's budget balance.

The recent dominant strand of the literature on wealth distribution has its emphasis on calibrated models of dynamic equilibrium economies with heterogeneous earnings and aptitudes.[9] Notably, Castaneda-Diaz Gimenez-Rios Rull (2003) calibrate a skewed

---

[8]See Mandelbrot (1960) for the early introduction of more general power laws and stochastic processes to obtain Pareto-Levy distributions; see also Reed-Jorgensen (2003) for Double Pareto-Lognormal distributions. Also, Nirei-Souma (2004) study multiplicative wealth accumulation models with stochastic rates of return and a reflective lower barrier (Kesten processes); Levy (2003) studies the implications of differential rate of return across groups; Solomon (1999) and Malcai et. al. (2002) study similar processes in which the rate of return on wealth accumulation is interdependent across different groups of individuals (Generalized Lotka-Volterra models); Levy (2003) shows that different rates of returns across non-interdependent groups generate wealth distributions which are Pareto only in the tail. Also, Das-Yargaladda (2003) and Fujihara-Ohtsuki-Yamamoto (2004) study stochastic processes in which individuals randomly interact and exchange wealth, and Souma-Fujiwara-Aoyama (2001) add network effects to such random interactions.

[9]Pareto (1897, 1909) himself had emphasized the role of heterogeneity to explain distribution the distribution of wealth and income. He explicitly noted that an identical stochastic process for wealth

8

strochastic distribution of skills, and hence of earnings, to quantitatively account for the U.S. wealth distribution and Gini coefficient. Other studies exploit different elements of persistent heterogeneity in preferences in addition to a skewed distribution of earnings.[10] For instance, Krusell and Smith (1998) match the skewness of the US wealth distribution in a dynastic model by introducing persistent heterogeneity in the discount factors of the dynasties.[11] Assortative matching of attributes can also be exploited to obtain or exacerbate the skewness of the distribution of wealth, along the lines of Mandelbrot (1962), Becker (1973), and also Lucas (1978). Quadrini (2000), for instance, calibrates the earning distribution to match the skewness of the US wealth distribution, and adopts an assortative matching model to capture the role of entrepreneurship. Cagetti-De Nardi (2000, 2003) adopt a similar approach based on differences in entrepreneurial skills. De Nardi (2004), on the other hand, exploits a non-homogeneous bequest function in an OLG model to match the concentration of wealth in US data.[12] A highly informative

across agents will not induce the skewed wealth distribution that we observe in the data (see Pareto (1897), Note 1 to #962, p. 315-316). Pareto therefore introduced skewness into the distribution of talents or the endowments of agents, which are then filtered into a the wealth distribution through a stochastic process driven by random returns (1897, Notes to #962, p. 416). He then considered the distribution of endowments as an empirical question to be recovered by inverting the process that maps abilities to wealth. Pareto's methodology gave rise to an exchange with Edgeworth, who later formalized this inverse map by the "method of translation, " and was one of the first to theorize that normally distributed aptitudes would yield a log-normal distribution of incomes. For an account of this exchange and the literature that followed, and later involved the mathematicians Cantelli and Frechet, see Chipman (1976), section 4.5.

[10]A mechanism which produces skewed distributions of earnings proposed by Roy (1950) is the multiplicative composition of several randomly distributed factors (e.g., talent attributes) which gives rise to a log-normal distribution of wealth. We should also note that Mincer (1958) derives a log-normal distribution of earnings from a simple human capital choice model.

[11]Without additional features however calibrated infinitely lived or dynastic agent models with skewed endowments or heterogenous discounting in preferences have difficulty genering the fat left tails of wealth distribution. See Cagetti and De Nardi (2005).

[12]Becker and Tomes (1979, section VI) generate skewness in income distribution by introducing het-

survey of this literature is given by Cagetti and De Nardi (2005).

This literature provides many important insights towards understanding which combination of economic factors better matches the properties of the empirical wealth distribution. In contrast to this paper however, these models are too complex to attempt an analytical characterization of the wealth distribution and hence to isolate and study in closed form how different economic factors contribute to the determination of the stationary wealth distribution. Furthermore, while we focus on the inter-generational transmission of wealth, that is on bequests rather than earnings, and show that this factor is sufficient to generate a wealth distribution with a heavy upper tail, the specific quantitative properties of the wealth distribution in these models are instead closely related to the underlying assumed distribution and skewness of labor endowments and abilities that generate skewed earnings.

In a methodological sense, therefore, this paper is more related to Wang (2006), as both papers derive a closed form characterization of the distribution of wealth. But the focus is different: Wang (2006) models an economy with infinitely lived agents and studies the distribution of wealth induced by stochastic earning processes, while we model an overlapping generations economy and study the distribution of wealth induced by inter-generational transmission through bequests.[13]

Finally, methods similar to those used to study the distribution of wealth have been used to study other distributions with heavy upper tails. See for instance the the early contribution of Simon-Bonini (1958) studying the statistical properties of the distribution of firms by size. Recently, in this context, Luttmer (2004, 2005) has obtained power laws by the explicit modeling of the entry and exit decisions of firms. Relatedly, Gabaix (1999)

---

erogeneity in the "propensity to invest in children." Huggett (1996) also calibrates an OLG model, with involuntary bequests and an exogenous log normal distribution of labor endowments within cohorts that matches the U.S. Gini coefficient for wealth but not the fat tails of the distribution.

[13]See also Brown-Channing-Chiang (2006a,b), who generate power law distributions of wealth via exogenous bequest rules.

developed a model of the growth of cities which generates a power law distribution of their size (a particular power law with power -1, in fact, which goes under the name of Zipf's Law).

2 Wealth accumulation in an OLG economy with bequests

Consider the Overlapping Generation (OLG) economy in Yaari (1965) and Blanchard (1985). Each agent at time t has a probability of death $\pi(t) = pe^{-pt}$. All agents have identical standard momentary utility from consumption. Agents also care about the bequest they leave to their children. We assume for simplicity that agents have a single child. At any time t an agent allocates his wealth between an asset and an annuity. The asset pays a return r, gross of taxes. We assume r is an exogenous constant (productivity) parameter. With perfect capital markets, by no-arbitrage, the annuity pays therefore a return $p + r$, where p is the probability of death.[14]

Let $c(s,t)$ and $w(s,t)$ denote, respectively, consumption and wealth at t of an agent born at s. Each agent's momentary utility function, $u(c(s,t))$, satisfies the standard monotonicity and concavity assumptions. Let $\omega(s,t)$ denote the amount invested in the asset at time t by an agent born at s, with wealth $w(s,t)$. Therefore $w(s,t) - \omega(s,t)$ denotes the amount that an agents invests in the annuity. If the agent dies at time t the amount bequeathed is the whole amount invested in the asset, $\omega(s,t)$. In other words, the asset is in fact effectively a bequest account. Letting b denote the estate tax , the agent's child inherits $(1-b)\omega(s,t)$.

Parents have a preference for leaving bequests to their children. In particular, we assume "joy of giving" preferences for bequests: the parent's utility from bequests is $\chi\phi((1-b)\omega(s,t))$, where ϕ denotes an increasing bequest function. Note that we have as-

---

[14]In particular, life insurance can be obtained by a negative position on the annuity.

sumed that the argument of the parents' preferences for bequests is after-tax bequests. At the end of this section we discuss how our analysis can be simply re-interpreted to encompass the case of pure altruism on the part of the parents, that is, the case in which parents internalize their children's utility rather than having direct preferences for bequests.

An agent born at time s receives, at birth, initial wealth $w(s, s)$ This initial wealth includes bequests and government transfers. It also contains the agent's expected present discounted value of lifetime earnings.[15] Importantly, the only stochastic component affecting wealth accumulation, in this economy, is the time of death and markets are complete.

We let τ denote the capital income tax, for simplicity imposed on the holdings of both the asset and the annuity.

The maximization problem of an agent born at time s involves choosing consumption and bequests paths, $c(s, t), \omega(s, t)$, to maximize

$$\int_t^\infty e^{(\theta+p)(t-v)} \left(u\left(c\left(s, v\right)\right) + p\phi\left((1-b)\omega\left(s, v\right)\right)\right) dv \tag{1}$$

subject to:

$$w\left(s, t\right) = w\left(s, s\right) + \int_s^t \left(\left(r + p - \tau\right) w\left(s, v\right) - p\omega\left(s, v\right) - c\left(s, v\right)\right) dv \tag{2}$$

and the transversality condition,

$$0 = \lim_{v \to \infty} e^{-\int_t^v r+p} w\left(s, v\right) dv.$$

In the interest of closed form solutions we assume

$$u(c) = \ln(c), \quad \phi(\omega) = \chi \ln \omega$$

The characterization of the optimal consumption-savings path is then straightforward.[16]

Proposition 1 *The consumption-savings path which solves the agent's maximization problem (1) is characterized by:*

$$c = \eta w, \quad \omega = \chi \eta w, \tag{3}$$

with $\eta = \frac{(p+\theta)}{p\chi+1}$ and

$$\dot{w}(s,t) = (r - \theta - \tau) \, w(s,t) \tag{4}$$

Notably, the growth rate of an agent's wealth, $g = r - \theta - \tau$, is independent of the preference parameter for bequests χ. A relatively low preference for bequests χ increases consumption as fraction of wealth but has no effect on the rate of growth of agent's wealth g. As a consequence, g decreases with the capital income tax τ but is independent of estate taxes b.

Pure Altruism. While we have solved for the agents' consumption-savings problem under the assumption of "joy of giving" preferences for bequests, the same analysis can be extended to the case of purely altruistic preferences. Consider to this effect the case of an agent who values his son's utility $\alpha \leq 1$.[17] The altruistic agent's maximization problem in recursive form is:

$$V\left(w(s,t)\right) = max_{c,\omega} \int_t^\infty e^{(\theta+p)(t-v)} \left(\ln c\left(s,v\right) + p\alpha V\left((1-b)\,\omega(s,v)\right)\right) dv$$

---

[16] We restrict parameters so that interior solutions obtain. We assume also that $r > -p$ to guarantee that the transversality condition is satisfied.

[17] Note that our formulation of altruism and intergenerational preferences is different from Phelps-Pollak (1968)'s inasmuch as it induces a time-consistent preference ordering over consumption sequences even for $\alpha < 1$.

subject to

$$\frac{dw\,(s,t)}{dt} = (r + p - \tau)\,w\,(s,t) - p\omega - c\,(s,t)$$

and the transversality condition. We show in Appendix A that optimal consumption-savings decisions of an altruistic agent correspond to those of an agent with "joy of giving" preferences with preference for bequest χ determined endogenously and equal to $\frac{1}{\theta+p(1-\alpha)}$. Therefore, for an altruistic agent,

$$c = (\theta + p\,(1 - \alpha))\,w, \quad \omega = \alpha w$$

and

$$\dot{w}(s,t) = (r - \theta - \tau)\,w(s,t)$$

Note that, when $\alpha = 1$ and the parent cares about his son as for himself, all of the wealth is deposited in the bequest account, that is, $\omega = w$, and it is fully inherited.

2.1 The aggregate economy

Regarding the demographics of the economy, we assume that the population is stationary: for any agent who dies at any time t there is a new agent born. Since each agent in the economy dies with probability p, at any time t, p agents die and the size of the cohort born at s is $pe^{-p(t-s)}$. The total population of the economy at any time t is therefore $\int_{-\infty}^{t} pe^{(s-t)p}ds = e^{(s-t)p}\mid_{-\infty}^{t} = 1$.

We also assume that, of the p agents dying at any time t, only $q < p$ leave an inheritance; $p - q$ die with no estate, e.g., because they have no preferences for bequests, $\chi = 0.$[18] Recall that, in Proposition 1, we have shown that the growth of individual

---

[18]This assumption is necessary in certain specifications of our economy to maintain a fraction of population with low wealth so as to keep the support of the wealth distribution sufficiently stationary; see also footnote 24. Alternatively, but equivalently, we could have postulated a constant inflow to the population at low wealth, e.g., of migrants.

wealth g is independent of χ. Agents who have a preference for bequest $(\chi > 0)$ consume a smaller fraction of wealth than agents who do not $(\chi = 0)$, but grow at the same rate g.

Let the aggregate economy's growth rate of wealth be denoted g'. Aggregate wealth is defined as:

$$W(t) = \int_{\infty}^{t} w(s,t) p e^{p(s-t)} ds$$

Let $W(s,t)$ denote the aggregate wealth at time t of all agents born at time s. Then,

$$\dot{W}(t) = W(t,t) - pW(t) + \int_{-\infty}^{t} \frac{dw(s,t)}{dt} p e^{p(s-t)} ds$$

Since the individual growth rate of wealth is constant across all agents in our economy,[19] $\frac{dW(s,t)}{dt} = (r - \theta - \tau) W(s,t)$ and

$$\dot{W}(t) = W(t,t) - pW(t) + (r - \tau - \theta) W(t) \tag{5}$$

where $W(t,t)$ is the initial wealth of all newborn agents

The growth rate of $W(t)$ is determined once we specify the initial wealth of all newborn agents at each time t, $W(t,t)$. In our economy $W(t,t)$ is composed of *i)* the financial wealth inherited from parents, *ii)* subsidies from the government, and possibly *iii)* the expected present discounted value of lifetime labor earnings.

Assuming government budget balance, subsidies must equal total tax revenues minus government expenditures. Suppose that a fraction γ of wealth constitutes government expenditures which are not re-distributed to agents in the economy.[20] Also, let $\mu = \frac{w - \omega}{w}$ denote the constant fraction of wealth, characterized in Proposition 1, agents with preferences for bequests $(\chi > 0)$ invest in the annuity. It follows then that aggregate inherited wealth in the economy is $q(1-\mu)(1-b)W(t)$ and that tax revenues net of expenditures is $q(1-\mu)bW(t) + \tau W(t) - \gamma W(t)$. Suppose also that the expected present

---

[19] Recall that, by Proposition 1, the growth rate g is independent of preferences for bequests, χ.

[20] Alternatively, but equivalently, we could assume that government expenditures finance the provision of a public good which enters additively separably into agents' preferences.

discounted value of lifetime labor earnings of an agent born at time t, let it be denoted x, is a constant fraction λ of wealth, $x = \lambda_x W$, so that it grows at the aggregate growth rate of wealth.

Since the aggregate wealth of newborn at t, $W(t, t)$ is comprised of aggregate inherited wealth, of tax revenues net of expenditures, and of earnings,

$$W(t, t) = (q(1 - \mu) + \tau - \gamma + p\lambda_x) W(t)$$

The dynamics of aggregate wealth then is

$$\dot{W}(t) = (r - \tau - \theta - p) W(t) + q(1 - \mu) W(t) + \tau W(t) - \gamma W(t) + p\lambda_x W(t)$$

and the growth rate of aggregate wealth in the economy is

$$g' = r - \theta - p + q(1 - \mu) - \gamma + p\lambda_x \tag{6}$$

The growth rate of aggregate wealth g' decreases with government expenditures as a fraction of wealth, γ. Importantly, g' decreases also with $p - q$, the density of agents which at any time t die with no bequests, and with $q\mu$, the fraction of the wealth of the agents which die at any period t which is annuatized, that is, not bequeathable. In other words, a high fraction of agents with no preference for bequests and/or a low fraction of bequeathable wealth (due in the model to low preferences for bequests) imply that a higher fraction of aggregate wealth is consumed and hence a lower aggregate growth rate. On the other hand, the growth rate of individual wealth g, as we noted, is independent of the preference for bequests of the agent, χ, and hence of μ. Also, the growth rate of aggregate wealth increases with the present value of earnings component λ_x, while the growth rate of individual wealth is independent of λ. The difference in the growth rate of individual and aggregate wealth is

$$g - g' = p - q(1 - \mu) - \tau + \gamma - p\lambda_x$$

In all of our subsequent analysis, the parameters of distribution of wealth and the expression for aggregate welfare will depend on $g - g'$, and in particular on the term $\gamma - \tau$,

and not on γ or τ separately. Therefore from here onwards, without loss of generality, we set $\gamma = 0$, and interpret τ as the tax rate on wealth net of the government expenditure rate γ.

2.2   Welfare policy

We assume tax revenues net of expenditures are positive, that is, $q(1-\mu)b + \tau > 0$. Government fiscal policy includes therefore a re-distributive component, in the form of a welfare policy. Also, note that our assumption that the expected present discounted value of earnings of agents born at time t, $x(t)$, is a constant fraction λ_x of aggregate wealth $W(t)$, implies that $x(t) = xe^{g't}$, for some $x > 0$.

The class of welfare policies we study guarantees that all agents born at any time t with earnings but no inheritance receive a transfer of wealth to bring them to a minimum wealth level $\underline{w}(t)$ which grows at the aggregate economy's rate g', that is, $\underline{w}(t) = \underline{w}e^{g't}$. In particular, we study *means-tested subsidies*:[21] all agents born at any t with initial wealth (inheritance plus earnings) less than $\underline{w}(t)$ get a transfer of wealth to bring them to $\underline{w}(t)$.[22]

The total amount of subsidies paid by the government in the form of means-tested subsidies at any time t depends on the distribution of wealth at t. In particular, the welfare policy subsidizes the wealth of those newborn whose parents are relatively poor

---

[21] See Appendix E for a discussion on the extension to *lump-sum subsidies*.

[22] Parents with "joy of giving" preferences for bequests bequeath amounts which are independent of these welfare transfers. Purely altruistic parents, in presence of such welfare transfers, would however choose to bequeath nothing at least until their own wealth at death is greater than $\frac{w(t)-x(t)}{(1-b)(1-\mu)}$, the amount of wealth at death with implies a bequest equal to $c(t) - x(t)$. With altruistic parents and means-tested subsidies, therefore, the optimal consumption savings path derived in Proposition 1 will be exact only for $w(t) \geq \hat{w}(t) \geq \underline{w}(t)$ where $\hat{w}(t) \rightarrow \underline{w}(t)$ as $\underline{w}(t) \rightarrow 0$: see the proof of Proposition 1 in Appendix A for details. However, it would be straightforward to design a non-linear scheme of welfare transfers with the property that the consumption savings path derived in Proposition 1 represents in fact the optimal path also with purely altruistic parents.

at death, that is, have wealth between $\underline{w}(t) - x(t)$ and the amount which implies a bequest equal to $\underline{w}(t) - x(t)$, $\frac{w(t)-x(t)}{(1-b)(1-\mu)}$.[23] Let $f(w,t)$ denote the distribution of wealth at time t. Total subsidies at time t are:

$$(p-q)\left(\underline{w}(t) - x(t)\right) + q \int_{(\underline{w}(t)-x(t))}^{((1-b)(1-\mu))^{-1}(\underline{w}(t)-x(t))} \left(\underline{w}(t) - x(t) - (1-b)(1-\mu)w\right) f(w,t)dw$$

Our analysis is simplified and results are sharper if we assume that the expected present discounted value of earnings, $x(t)$, is zero, that is, if we concentrate on the intergenerational transmission of wealth. We proceed by imposing this assumption in the next section. In this case total subsidies are:

$$(p - q)\underline{w}(t) + q \int_{\underline{w}(t)}^{((1-b)(1-\mu))^{-1}\underline{w}(t)} \left(\underline{w}(t) - (1-b)(1-\mu)w\right) f(w,t)dw \tag{7}$$

However, at the end of the section (see Proposition 6)we will turn back to study the economy with earnings to show that our results hold unchanged for the upper tail of the wealth distribution.

3   The distribution of wealth

We study the dynamics of the distribution of wealth in economy with inheritance and estate taxes introduced in the previous section. We solve for both the transitional dynamics and the stationary distribution. We study conditions under which the stationary distribution is Pareto.

The dynamics of the distribution of wealth $f(w,t)$are described by a linear partial differential equation (PDE) with variable coefficients, an initial condition for the initial

---

[23]Note that such subsidies can be supported by a stationary tax policy (with constant rates τ, b, as we have assumed) only if the distribution of wealth is stationary (independent of t) or if we allow the government to run fiscal deficits and surpluses and only require a balanced budget inter-temporally, rather than for all t.

wealth distribution, and a boundary condition that reflects the injection of wealth to newborns under our welfare policies.

We restrict parameters so that individual wealth accumulates faster than aggregate wealth, that is:

$$g - g' = p - q(1 - \mu) - \tau > 0 \tag{8}$$

We will show later that this condition avoids a degenerate wealth distribution.

Let $\sigma(w) = \frac{w}{(1-b)(1-\mu)}$ denote the wealth a parent needs to have at time of death t for his heir born at $t + \Delta$ to inherit wealth w.

The PDE describing the evolution of the distribution of wealth is obtained as the Chapman-Kolmogorov equation which governs the dynamics of $f(w,t)$ (its derivation is detailed in Appendix A):

$$\frac{\partial f(w,t)}{\partial t} = -(p+g) f(w,t) + q \frac{\partial \sigma(w)}{\partial w} f(\sigma(w),t)) - gw \frac{f(w,t)}{\partial w} \tag{9}$$

At time 0 the distribution of wealth $w \in (\underline{w}, \infty)$ is exogenous. Let it be denoted $h(w)$. We assume for simplicity that at time $t = 0$ all agents have wealth greater than minimal wealth:

$$h(w) = 0 \text{ for any } w \leq \underline{w}$$

The initial condition of the PDE is then:

$$f(w,0) = h(w) \tag{10}$$

The distribution of wealth at time t must also satisfy the boundary condition (derived in Appendix A):

$$f(\underline{w}(t), t) = \frac{p - q}{g} \frac{1}{\underline{w}(t)} + \frac{1}{g} \frac{1}{\underline{w}(t)} q \int_{\underline{w}(t)}^{\sigma(\underline{w}(t))} f(w,t) dw \tag{11}$$

This boundary condition guarantees that, at each t, the population size is constant and normalized to 1; that is, $\int f(w,t) dw = 1$. Note that $f(\underline{w}(t), t)$, the density of wealth at

$w = \underline{w}(t)$, is composed of the density of wealth corresponding to the $p - q$ agents who do not receive any inheritance, $\frac{p-q}{g-g'}\frac{1}{\underline{w}}$, and of the the agents whose inheritance at t is below $\underline{w}(t)$, $\frac{1}{g}\frac{1}{\underline{w}}q\int_{\underline{w}(t)}^{\sigma(\underline{w}(t))} f(w,t)dw$.

Formally, our problem is the following: find a density $f(w,t)$ which satisfies the PDE (9) for all $w > \underline{w}(t)$, the initial condition (10), and the boundary condition (11).The mathematical problem is non-standard inasmuch as *i)* in the PDE, the unknown density f is evaluated at different arguments, w and $\sigma(w)$ and *ii)* the boundary condition is not independent of the unknown density f.

It will be convenient to work in variables discounted by the aggregate economy's growth rate g'. For this purpose define discounted wealth $z = we^{-g't}$. Note that the support of z is stationary and equal to $(\underline{w}, \infty)$. The PDE which we obtain after the necessary transformations for discounted variables is:

$$\frac{\partial f(z,t)}{\partial t} = -(p + g - g') f(z,t) + q\frac{\partial \sigma(z)}{\partial z} f(\sigma(z),t)) - (g - g') z \frac{f(z,t)}{\partial w} \tag{12}$$

with initial condition:

$$f(z,0) = h(z) \tag{13}$$

and boundary condition:

$$f(\underline{w},t) = \frac{p-q}{g-g'}\frac{1}{\underline{w}} + \frac{1}{g-g'}\frac{1}{\underline{w}(t)}q\int_{\underline{w}}^{\sigma(\underline{w})} f(z,t)dz \tag{14}$$

To solve (9) under (13) and (14) we apply the "method of characteristics" as detailed in Appendix C.

Lemma 1 *There exists a distribution of discounted wealth $f(z,t)$ which satisfies (12) as well as (13). It is characterized by:*

$$f(z,t) =$$

$$
\begin{cases}
\left(\frac{z}{\underline{w}}\right)^{-\frac{p}{g-g'}-1} f\left(\underline{w}, t - \tau\left(z, \underline{w}\right)\right) + \\
\qquad + q \int_{\underline{w}}^{z} \frac{\partial \sigma(y)}{\partial y} f\left(\sigma(y), t - \tau\left(z, y\right)\right) (y)^{\left(\frac{p}{g-g'}\right)} (g - g')^{-1} (z)^{-\left(\frac{p}{g-g'}+1\right)} dy & \text{for } z \in \left(\underline{w}, \underline{w} e^{(g-g')t}\right) \\
\\
e^{-(p+g-g')t} h\left(z e^{(g-g')t}\right) \\
\qquad + q \int_{\underline{w}}^{z} \frac{\partial \sigma(y)}{\partial y} f\left(\sigma(y), t - \tau\left(z, y\right)\right) (y)^{\left(\frac{p}{g-g'}\right)} (g - g')^{-1} (z)^{-\left(\frac{p}{g-g'}+1\right)} dy & \text{for } z \geq \underline{w} e^{(g-g')t}
\end{cases}
$$

$$(15)$$

where $\tau(z,y) = \frac{\ln z}{\ln y} \frac{1}{g-g'}$

This characterization has an interesting economic interpretation. Notice that $\tau(z,y) = \frac{\ln z}{\ln y} \frac{1}{g-g'}$ represents the age of an agent who has wealth z at time t and was born with wealth y. The age of an agent who has wealth z at time t and was born with wealth $\underline{w}$ is then $\tau(z, \underline{w})$. Consider the density of any discounted wealth level $z \in \left(\underline{w}, \underline{w} e^{(g-g')t}\right)$. The first component of the density $f(z,t)$ in (15) is $\left(\frac{z}{\underline{w}}\right)^{-\frac{p}{g-g'}-1} f\left(\underline{w}, t - \tau\left(z, \underline{w}\right)\right)$. It repre-sents the density of agents who have entered the economy with wealth $\underline{w}$, have never died since, and have reached wealth z at t. It is determined by the boundary condition at time $t - \tau(z, \underline{w})$. Similarly, the second component of the density $f(z,t)$ in (15) is $q \int_{\underline{w}}^{z} \frac{\partial \sigma(y)}{\partial y} f\left(\sigma(y), t - \tau\left(z, y\right)\right) (y)^{\left(\frac{p}{g-g'}\right)} (g - g')^{-1} (z)^{-\left(\frac{p}{g-g'}+1\right)} dy$. It represents the den-sity of agents who have entered the economy with some wealth y, have never died since, and have reached wealth z at t. Consider finally the density of discounted wealth levels z at time t greater than $\underline{w} e^{(g-g')t}$. The only agents who can possess such a discounted wealth level are: *i)* those agents who were born at time 0 and have never died, *ii)* the children of those agents who have died at some time $t' < t$ and left inheritance larger than $\underline{w} e^{(g-g')t'}$. The density of these agents is represented by the second line of (15).

The distribution of wealth $f(z,t)$ must then satisfy (15) as well as (14). It is in general impossible to find a closed form solution unless the boundary condition (14) has

the property that $f(\underline{w}, t)$ is constant in t. We will discuss two a special economies for which this is the case in Section 3.1.

We can nonetheless study the limit distribution of the dynamics of $f(z, t)$. First of all we can show (see the proof of Lemma 2 in Appendix A) that the density of discounted wealth levels z at time t which are greater than $\underline{w}e^{(g-g')t}$, represented by the second line of (15), declines with time. It is in fact bounded above by $e^{-(p-q+g-g')t}h\left(ze^{-(g-g')t}\right)$. It therefore declines at a rate (greater than) $p - q + g - g'$ and vanishes for $t \to \infty$.

Lemma 2 *The distribution of wealth $f(z, t)$ which satisfies (15) as well as (14) has a stationary distribution, $f(z)$, which solves the following integral equation:*

$$f(z) = \left(\frac{z}{\underline{w}}\right)^{-\left(\frac{p}{g-g'}+1\right)} f(\underline{w}) + q \int_{\underline{w}}^{z} \frac{\partial \sigma(y)}{\partial y} f\left(\sigma(y)\right)(y)^{\left(\frac{p}{g-g'}\right)} (g-g')^{-1} (z)^{-\left(\frac{p}{g-g'}+1\right)} dy \tag{16}$$

for

$$f(\underline{w}) = \frac{p-q}{g-g'}\frac{1}{\underline{w}} + \frac{1}{g-g'}\frac{1}{\underline{w}}q \int_{\underline{w}}^{\sigma(\underline{w})} f(z)dz. \tag{17}$$

3.1 Pareto distributions

The integral equation (16) can be solved for quite generally. We proceed however by studying first two special simple economies which illustrate the inter-generational transmission mechanism we have modeled. The two special economies we study are characterized by extreme and opposite behavior in terms of bequests, full inheritance with no estate taxes and no inheritance, but nonetheless both display a stationary distribution of wealth which is Pareto.

Full inheritance We first study the an economy in which agents leave all of their wealth as inheritance to their children. This requires a large enough χ as well as no estate taxes $(b = 0)$.

Recall however that at each time t $p-q$ agents die without leaving bequests and hence $p-q$ agents are born with minimal wealth $\underline{w}$. If $\mu = 0, x = 0$, it follows immediately that the boundary condition (14) requires:

$$f\left(\underline{w}, t\right) = \frac{p-q}{g-g'} \frac{1}{\underline{w}},$$

Furthermore, from (8), $g - g' = p - q - \tau$.

We are then ready to characterize the dynamics of the distribution of wealth in this economy.

Proposition 2 *The economy with full inheritance and no estate taxes has the following distribution of discounted wealth at each time t:*

$$f(z,t) \begin{cases} \frac{p-q}{p-q-\tau} \ \underline{w}^{\frac{p-q}{p-q-\tau}} \ z^{-\left(\frac{p-q}{p-q-\tau}+1\right)} & \text{for } z \in \left(\underline{w}, \underline{w}e^{(p-q-\tau)t}\right) \\ e^{-(p+p-\tau)t} h\left(ze^{-(p-q-\tau)t}\right) & \text{for } z \geq \underline{w}e^{(p-q-\tau)t} \end{cases} \tag{18}$$

It is a truncated Pareto distribution in the range $\left(\underline{w}, \underline{w}e^{(p-q-\tau)t}\right)$. The ergodic distribution of discounted wealth is

$$f(z) = \frac{p-q}{p-q-\tau} \ \underline{w}^{\frac{p-q}{p-q-\tau}} \ z^{-\left(\frac{p-q}{p-q-\tau}+1\right)}$$

which is a Pareto distribution with exponent $P = \frac{p-q}{p-q-\tau}$ and finite mean.[24]

The stationary distribution of wealth in this economy is characterized by the single parameter P. Wealth inequality, for instance, is inversely related to the Pareto exponent P. Let in fact G denote the Gini coefficient of the stationary distribution of wealth, a standard measure of wealth inequality. For a Pareto distribution it is well known that G is related to P by the expression:[25]

$$G = \frac{1}{2P-1}$$

---

[24]Note that the stationary distribution of wealth is not a Pareto distribution in this economy if $q = p$, that is, if all agents leave full inheritance. In this case, the initial distribution of discounted wealth $h(z)$ remains unchanged over time.

[25]See e.g., Chipman (1974).

Several properties of $P = \frac{p-q}{p-q-\tau}$ in this economy are worth noticing. First of all, P de-
pends positively on the capital tax τ, and hence the capital tax reduces wealth inequality.
(Note that τ, in fact, represents capital income taxes net of non-redistristributive gov-
ernment expenditures and hence is a measure of the re-distributional component of fiscal
policy.) Furthermore when capital taxes τ tends to 0, P tends to 1 (Zipf's Law)), the
Pareto distribution looses its mean and it is maximally unequal (the Gini coefficient tends
to 1). Finally, P depends positively on rate at which agents die with no inheritance,
the numerator $p - q$, and negatively on the difference between the individual and the
aggregate growth rate of wealth, the denominator $g - g' = p - q - \tau$. Intuitively, in fact,
a high rate at which agents die with no inheritance tends to dissipate wealth, compress
its distribution, and hence to limit inequality (the children of the agents who die with no
inheritance also receive redistributive welfare subsidies), while a high difference between
the individual and the aggregate growth rate of wealth tends to spread the distribution
and hence to increase inequality. The rate at which agents die without an inheritance,
$p - q$, has hence both a negative effect on inequality (through the numerator of P, as
observed) as well as a positive effect, by increasing the difference between the individual
and the aggregate growth rate of wealth in our economy (through the denominator of P).
The positive effect is a consequence of the fact that the rate at which agents die without
an inheritance affects negatively aggregate consumption and affects positively aggregate
savings and hence the growth rate of aggregate wealth g'. It is however straightforward
to see that, for positive capital taxes τ, the composite effect of the rate at which agents
die without an inheritance on wealth inequality, $p-q$, is positive (that is, that the growth
effect in the denominator of P dominates).

No inheritance   Consider now another special economy, in fact one with opposite
behavior in terms of bequests, in which agents only invest in annuities and leave no
bequests, $\mu = 1$. This is the case, for instance, if all agents have no preferences for

bequests ($\chi = 0$). It would also the case if bequest declined with estate taxes[26] and estate taxes were expropriatory ($b = 1$).

In this economy, all p newborns at time t receive $\underline{w}$. Consequently, the density of wealth at the boundary $\underline{w}$ is constant over time and the boundary condition is reduced to:

$$f(\underline{w}, t) = \frac{p}{g - g'} \frac{1}{\underline{w}}, \tag{19}$$

while the initial condition is the same as in (13). Furthermore, from (8), $g - g' = p - \tau$.

It is possible then to characterize the dynamics of the distribution of wealth in this economy.

Proposition 3 *The economy without bequests has the following distribution of discounted wealth at each time t:*

$$f(z, t) = \begin{cases} \frac{p}{p - \tau} \underline{w}^{\frac{p}{p-\tau}} z^{-\left(\frac{p}{p-\tau} + 1\right)} & \textit{for } z \in \left(\underline{w}, \underline{w} e^{(p-\tau)t}\right) \\ e^{-(p+p-\tau)t} h\left(z e^{-(p-\tau)}\right) & \textit{for } z \geq \underline{w} e^{(p-\tau)t} \end{cases} \tag{20}$$

$f(z, t)$ is a truncated Pareto distribution in the range $\left(\underline{w}, \underline{w} e^{(p-\tau)t}\right)$. The ergodic distribution of discounted wealth is

$$f(z) = \frac{p}{p - \tau} \underline{w}^{\frac{p}{p-\tau}} z^{-\left(\frac{p}{p-\tau} + 1\right)}$$

which is a Pareto distribution with exponent $P = \frac{p}{p-\tau}$ and finite mean.

Both the economy with no inheritance and the economy with full inheritance have a stationary distribution of discounted wealth that is Pareto. But do both of these economies, despite their difference with respect to bequests, nonetheless have a Pareto distribution of discounted wealth at steady state? First of all note that the economy with full inheritance is homeomorphic to an economy without no inheritance in which agents die with probability $p - q$. A sequence of generations that pass on their wealth (a dynasty)

---

[26]This is not the case in our model because we assumed logarithmic preferences for bequests.

is in fact the natural unit of analysis in the full inheritance economy, corresponding to what an agent is in the no inheritance economy, and such dynasties are broken (die) only with probability $p - q$.

Furthermore, notice that the stochastic process generating the distribution of discounted wealth in these economies has a simple character: for each agent discounted wealth grows exponentially until a Poisson distributed stopping time hits, when discounted wealth drops to a lower bound. This class of stochastic processes is studied formally already by Cantelli (1921) and then by Fermi (1949), and it is known to aggregate into a Pareto distribution at steady state.[27]

Also, the Pareto exponent P has a related characterization in both economies. In summary, the Pareto distribution results as a consequence of the balancing of two opposite forces, wealth accumulation and redistribution. In these simple economies these forces take the form, respectively, of *i)* the growth of individual wealth relative to aggregate wealth, which tends to spread the distribution, and *ii)* death and redistributive welfare, which tends instead to compress the distribution.

The general case We are now ready to study the case in which in which agents leave part of their wealth as inheritance to their children and estate taxes are imposed; that is, the case in which $0 < \mu, b < 1$. While the analysis of the distribution of wealth in this economy is more involved, we can nonetheless show that the stationary distribution of discounted wealth remains Pareto and we can characterize its properties essentially in close form.

We study directly the stationary distribution as in this case we cannot analytically solve (15) for the transitional dynamics of $f(z, t)$. We therefore look for a function $f(z)$ which satisfies the integral equation (16) and the boundary condition (17).

---

[27]Various stochastic processes for individual wealth are known to aggregate into a Pareto distribution of wealth in the population; see Sornette (2000) for a technical review and Chipman (1976) for a careful and outstanding account of the historical contributions of this subject; see also Levy (2003).

We use the transformation $j = \sigma(y) = \frac{y}{(1-\mu)(1-b)}$ and obtain, from (16):

$$
\begin{aligned}
f(z) &= \left(\frac{z}{\underline{w}}\right)^{-\left(\frac{p}{g-g'}+1\right)} f(\underline{w}) \\
&\quad + q(g-g')^{-1} \int_{\frac{\underline{w}}{(1-\mu)(1-b)}}^{\frac{z}{(1-\mu)(1-b)}} f(j) \left[((1-\mu)(1-b)j)^{\left(\frac{p}{g-g'}\right)} z^{-\left(\frac{p}{g-g'}+1\right)}\right] dj
\end{aligned}
\tag{21}
$$

Recall that, from (8), $g - g' = p - q(1-\mu) - \tau$ We proceed by guessing a Pareto distribution for $f(z)$:

$$
f(z) = \frac{p - aq(1-\mu)(1-b)}{p - q(1-\mu) - \tau} \, \underline{w}^{\frac{p-aq(1-\mu)(1-b)}{p-q(1-\mu)-\tau}} \, z^{-\left(\frac{p-aq(1-\mu)(1-b)}{p-q(1-\mu)-\tau}+1\right)}
\tag{22}
$$

and then solve for the parameters a to satisfy, respectively, (21) and the boundary condition (17).

After some algebra, we can show that the guess (22) satisfies (21) if and only if a solves the fixed point equation:

$$
a = ((1-\mu)(1-b))^{\left(\frac{p-aq(1-\mu)(1-b)}{p-q(1-\mu)-\tau}-1\right)}
\tag{23}
$$

It is straightforward to show that (23) has a unique fixed point, which we denote a^*, and that $0 < a^* < 1$. The boundary condition (17) is also satisfied and $\int_{\underline{w}}^{\infty} f(z)dz = 1$. We summarize this analysis with the following result.

Proposition 4 *The economy with inheritance, estate taxes, and means-tested subsidies has a stationary distribution of discounted wealth*

$$
f(z) = \frac{p-a^*q(1-\mu)(1-b)}{p-q(1-\mu)-\tau} \, \underline{w}^{\frac{p-a^*q(1-\mu)(1-b)}{p-q(1-\mu)-\tau}} \, z^{-\left(\frac{p-a^*q(1-\mu)(1-b)}{p-q(1-\mu)-\tau}+1\right)},
\tag{24}
$$

for $0 < a^* < 1$ satisfying (23)

which is a Pareto distribution with exponent $P = \frac{p-a^*q(1-\mu)(1-b)}{p-q(1-\mu)-\tau}$ and finite mean. Furthermore, $f(z)$ is ergodic.

Furthermore it is straightforward to show, following the analysis of the special economies we have studied previously, that the Gini coefficient for this economy is in between that of the full inheritance economy and that of no inheritance economy.

In the general economy, as in the simple economies studied above, the Pareto distribution can be usefully interpreted as resulting from the interplay of wealth accumulation and redistribution. The growth rate of discounted individual wealth is $g - g' = p - q(1 - \mu) - \tau$, the denominator of the exponent P. As a consequence, wealth inequality decreases with the aggregate growth effects of the re-distributional component of fiscal policy, $q(1 - \mu) - \tau$. It decreases also with the density of agents who receive the welfare subsidies at birth, which can be shown to be equal to $p - a^* q(1 - \mu)(1 - b)$, the numerator of P.

We study in detail the dependence of the Pareto exponent and of the Gini coefficient of the stationary distribution of wealth in the next section. Wealth inequality, however, also depends on the strength of the bequest motive. An increase of the preference for bequest, χ, or of the fraction of agents with such preference, q, increases the fraction of total wealth left as inheritance, $q(1 - \mu)$. As a consequence, the aggregate growth rate of the economy increases without raising the growth rate of individual wealth, and the Pareto coefficient rises, decreasing wealth inequality.

As a corollary of Proposition 4 we may also characterize the distribution of wealth conditional on age. Let $A(z; n)$ characterize the stationary wealth distribution of those agents who have reached age n. An n year old with current wealth z must have started life with wealth $z e^{-(g-g')n} \geq \underline{w}$. Note that at the stationary distribution of wealth $f(z)$, given by Proposition 4, the inflow of the wealth distribution of newborns exactly offsets the outflow of the wealth distribution of the dying agents, so $f(z)$ remains invariant. Since the death rate is p the current wealth distribution of n-year old agents is given by $pf\left(z e^{-(g-g')n}\right) e^{-pn}$ for $z e^{-(g-g')n} \geq \underline{w}$.

Corollary 5 *The stationary wealth distribution of agents n years old is*

$$A(z; n) = pf\left(z e^{-(g-g')n}\right) e^{-pn} \qquad z \geq e^{(g-g')n}\underline{w}$$

where the density $f(z)$ is given by Proposition 4.

Since $f(z)$ is a Pareto distribution on $z \geq \underline{w}$, it follows that $A(z; n)$ is also a Pareto density on $z \geq \underline{w}e^{(g-g')n}$ scaled down by the factor pe^{-pn}. The Corollary above clearly illustrates the role of inheritance rather than age in generating the skewness of the wealth distribution within each age cohort in our economy.

As noted, our results in this section pertain to the distribution of wealth as generated by intergenerational transmission and redistributive fiscal policy. In other words, we have shut off an important component of wealth, earnings, by assuming $x(t) = 0$. How is our analysis changed if we allow for earnings, that is to say, if

$$x(t) = xe^{g't}$$

for some $x > 0$?[28]

The next proposition shows that the economy of our model with no earnings provides an approximation to the economy with earnings, in the sense that our results are unchanged as long as we interpret them as results about the tail of the wealth distribution rather than the whole distribution.[29]

Proposition 6 *The economy with inheritance, estate taxes, and earnings has a stationary distribution of discounted wealth with the following properties:*i) *for any z, it is bounded below by a Pareto distribution with exponent*

$$\frac{p}{p - q(1 - \mu) - \tau} \tag{25}$$

---

[28]Our analysis can be straightforwardly extended to account for heterogeneous earnings across agents. On the other hand, as previously noted, our analyis requires that we assume that financial markets allow agents to diversify all earnings uncertainty so that only expected present discounted values of earnings enters each agent's consumption-saving problem. Relaxing this assumption is beyond our technical abilities. Stochastic earning processes, however in a model with no intergenerational transmission, are studied by Wang (2006).

[29]The proof of the proposition below is in Appendix E.

and it is bounded above by a Pareto distribution with exponent

$$\frac{p - a^* q \left(1 - \mu\right) \left(1 - b\right)}{p - q \left(1 - \mu\right) - \tau} > 1, \quad \text{for } 0 < a^* < 1 \text{ satisfying (23)}$$

ii) *for large z, it is approximated by a Pareto distribution with exponent* $\frac{p - a^* q (1-\mu)(1-b)}{p - q(1-\mu) - \tau} >$
1, *for $0 < a^* < 1$ satisfying (23).*

In other words, the proposition shows that in our model it is the inter-generational bequest mechanism that determines the heavy upper tail of the wealth distribution. Of course how far in the tail the effect of earnings diminish significantly is a matter for empirical work. Piketty-Saez (2003) look at this question with time series data for top income shares (distinguishing capital income from wages and other forms of income) which they construct from individual fiscal tax returns.

4   Redistributive Policies

In this section we study the effects of changes in fiscal policy, that is, changes in estate taxes b and capital income taxes τ, on the stationary distribution of discounted wealth, as parametrized by its Pareto exponent and the minimal wealth which can be supported by welfare. Thanks to Proposition 6 the analysis is pursued for the economy without earnings but is directly applicable to the upper tail of the economy with earnings (the Pareto exponent is inversely related the heaviness of the tail, as measured, e.g., by top wealth shares). Furthermore, we characterize optimal redistributive taxes with respect to an utilitarian social welfare measure.

4.1   Positive effects of fiscal policies

Fiscal policies, in the form of changes in estate taxes b and capital income taxes net of non-redistributive government expenditures τ, have a direct effect on the Pareto exponent

of the distribution of discounted wealth. We have shown in fact in Proposition 4 that the stationary distribution of discounted wealth is a Pareto distribution with finite mean whose exponent is:

$$P = \frac{p - a^* q \left(1 - \mu\right) \left(1 - b\right)}{p - q \left(1 - \mu\right) - \tau}, \quad \text{with } a^* = \left(1 - \mu\right)^{\left(\frac{p - a^* q (1-\mu)(1-b)}{p - q(1-\mu) - \tau} - 1\right)} \tag{26}$$

Fiscal policies, therefore, directly affect wealth inequality as measured by the Gini coefficient of the distribution of wealth, since, as we noted $G = \frac{1}{2P-1}$. But, keeping government expenditures constant as a fraction of wealth, changes in estate taxes and capital income taxes are purely redistributive through welfare policy financing. Changes in estate taxes b and capital income taxes net of non-redistributive government expenditures τ therefore also indirectly affect the minimal wealth which can be supported by welfare, $\underline{w}$. More specifically, the government budget constraint at the stationary distribution (24) can be written as (see the derivation in Appendix A):

$$\underline{w} = \frac{\left(\tau + bq(1 - \mu)\right) M}{p - q(1 - \mu)(1 - b) \left(a^* + \frac{P}{P-1}(1 - a^*)\right)} \tag{27}$$

where a^* solves (23) and M denotes the discounted mean wealth which is independent of redistributive fiscal policy changes.[30]

We can then characterize the effects of changes in fiscal policy b and τ on the stationary distribution of discounted wealth.

Proposition 7 *The Gini coefficient of the economy's stationary distribution of discounted wealth is decreasing in capital income taxes net of non-redistributive government expenditures τ and is non-increasing in estate taxes b. Perfect equality $(G = 0, P = \infty)$ is attained for $\tau = p - q \left(1 - \mu\right)$, independently of b. Moreover, the minimal wealth which*

---

[30]It can be easily shown that equivalent formulations of the government budget constraint are:

$$\underline{w} = \frac{P - 1}{P} M = \frac{1 - 2G}{1 + 2G} \tag{28}$$

31

can be supported by welfare, $\underline{w}$, is increasing in τ and non-decreasing in b. Thus as $P \to \infty$, and $G \to 0$, perfect equality is reached when minimum wealth is equal to mean wealth: $\underline{w} = M$.

The effects of taxes on the Pareto exponent, and therefore on inequality, operate through several channels. To the extent that capital taxes slow the growth of individual wealth relative to the growth of aggregate wealth (the denominator in the expression for P), inequality decreases.

In addition, estate and capital taxes affect the numerator of the expression for P, $p - a^* q (1 - \mu) (1 - b)$, which has the interpretation of the fraction of the agents that inherit wealth below $\underline{w}$, and hence are supported by the welfare policy. Since higher taxes increases the number of people who need be subsidized, the net effect of taxes on inequality is not immediately clear by inspection, but the Proposition above proves that in fact capital and estate taxes reduce inequality. Note also that the effect of capital income taxes on P becomes dominant as τ becomes large. As τ rises towards its upper bound, $p - q (1 - \mu)$, the Pareto exponent becomes large and tends towards infinity. Consequently the Gini coefficient is reduced, and the wealth distribution becomes more equal. As the distribution becomes more highly peaked, the expression $a^* (1 - \mu) (1 - b) = ((1 - \mu) (1 - b))^P$, representing the fraction of the q agents that inherit wealth above $\underline{w}$, declines. Consequently, the effect of estate taxes b decline as well: with small a^* the effect of b on $P = \frac{p - a^* q (1 - \mu)(1 - b)}{p - q(1 - \mu) - \tau}$ becomes negligible. It follows that the higher is the value of τ, the more insignificant is the effect of the estate taxes b on the Pareto and Gini coefficients.[31]

---

[31] Interestingly, Castaneda-Diaz Gimenez-Rios Rull (1993) also find small effects of estate taxes on the distribution of wealth in an equilibrium economy where the distribution of earnings are calibrated to match the wealth distribution in the US.

4.1.1 Fiscal policy in a calibrated economy

In this section we provide a calibration of our economy with two objectives. First of all, we aim at better illustrating the effects of fiscal policies on wealth inequality. Furthermore, we attempt a first assessment of the relative importance of inter-generational transmission of wealth, *old-money*, in the determination of observed wealth inequality levels. In particular, we compute the Gini coefficient of the stationary distribution of the calibrated economy. While the Gini coefficient is computed for the economy with no earnings, it is positively related to the heaviness of the upper of the wealth distribution, and the tail, we have shown, is invariant to the introduction of the expected present discounted value of their earnings into the agents' initial wealth. Of course, allowing for a stochastic process for earnings, especially one with a skewed stationary distribution as in Castaneda-DiazGimenez-Rios Rull (2003), would add to the stationary wealth inequality of the economy. In this sense our calibration exercise is an attempt to estimate the fraction of the observed wealth inequality in the U.S.that can be attributed to the inter-generational wealth transmission mechanism.

The deep parameters of our economy consist of the probability of death p, the proportion of agents who leave bequests q, the discount rate θ, the preference for bequest parameter χ, and the interest rate r. We choose $p = .016$ for an expected productive life of $p^{-1} = 62$ years. To calibrate the stationary distribution of wealth, in fact, we only need to set $q(1 - \mu)$, rather than q and $1 - \mu$ (hence χ) independently. We then calibrate $\frac{q}{p}(1 - \mu)$ to match the proportion of non-annuitized wealth. Operationally defining annuitized wealth involves several conceptual complications. While the private annuity markets are thin, we consider social security, certain employee pension plans, and 401K retirement accounts that can be considered as annuities reserved for retirement.[32] More specifically, Auerbach et al. (1995) report data for the wealth

---

[32]These annuities are not necessarily voluntary but they can be undone in the market by purchasing life insurance, which are negative annuities. Furthermore inheritances are also diluted by charitable

composition of U.S. males and females from 20 to 89 years of age in 1990. We compute (from Tables 2a-b and 3a-b) the fraction of wealth held in non-annuitized assets in 1990 as $\frac{\text{Non Human Wealth}}{\text{Non Human Wealth + Private Pensions + Social Security}}$ and obtain 0.4 (and hence $q(1-\mu) = .4p = .0064$). We also choose a 4% annual discount rate θ and an 8% gross interest rate r.[33]

Figure 1, shows the relationship between the Gini coefficient of inequality G and the taxes b and τ for our calibrated economy. It is apparent that high capital income taxes are required to generate Gini coefficients above 0.6, while estate taxes have little effect per se; see Appendix F for the tabulated values of G as a function of taxes b and τ.

Figure 2, shows instead the relationship between the ratio of minimum to average wealth, $\frac{w}{M}$, and taxes b and τ.

To assess the relative importance of inter-generational transmission of wealth in the determination of observed wealth inequality levels we need a calibration of fiscal policy. Recent tax return data from the Internal Revenue Service show that in 2003 taxable estates faced an average effective tax rate of only 19% (see e.g., Friedman-Carlitz, 2005).[34] Setting $b = .19$, the imputed flow of bequests as a share of non-human wealth in the calibration is therefore $bq(1-\mu) = 0.001\,216$, corresponding to 0.36% of GDP .[35] Welfare

---

giving. Another complication is that bequests of married couple to children occur in steps since the surviving spouse inherts a fraction of the estate and the full bequest accrues to the children after the death of th surviving spouse. On the other hand bequest data does not include inter-vivos transfers.

[33]We have also experimented with various values for r : the Gini coeffient in our model is insensitive to changes in r.

[34]Under the 2002 tax code estates under $1 million are exempt, and tax rates are progressive up to a top rate of 50%. In fact in 2003 taxable estates between $5 and $10 millions faced the highest effective tax rates, about 29%; the largest estates, those over $20 millions faced only a 16.5% effective tax rate because of the size of their charitable bequests.

[35]In our calculations, the non-human wealth to GDP ratio is 3.At 8% return, in fact, non-human wealth produces capital income that corresponds to a twelveth of itself, and a quarter of GDP, with the rest of GDP coming from earnings and other sources.
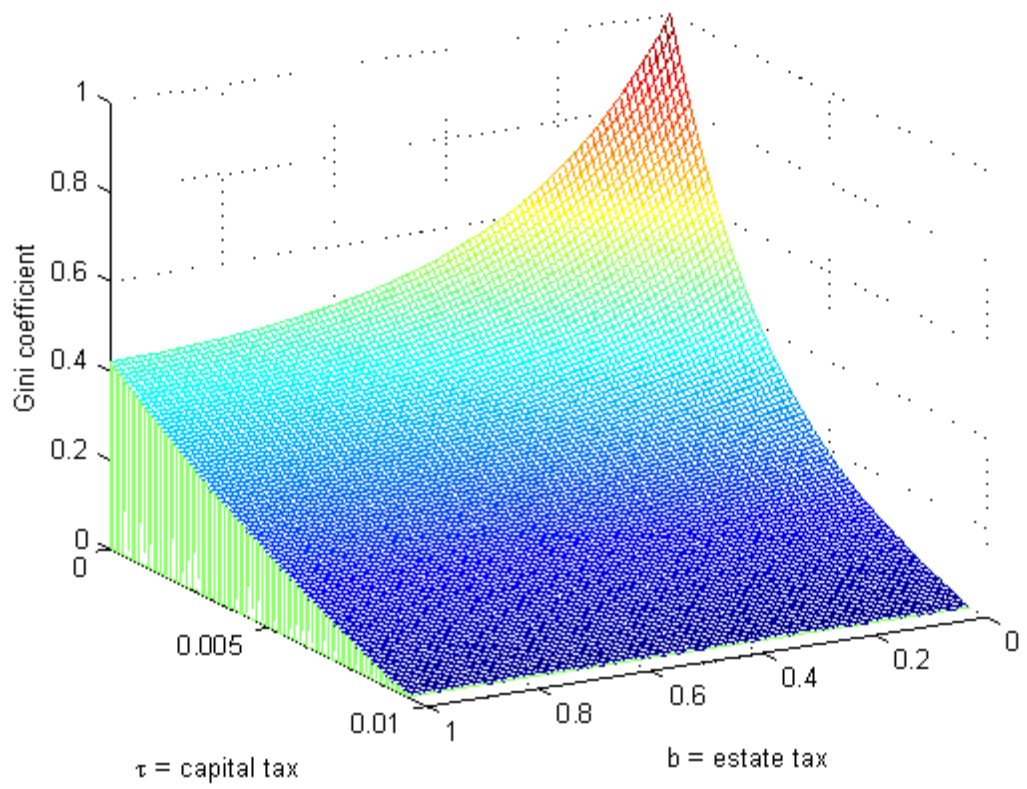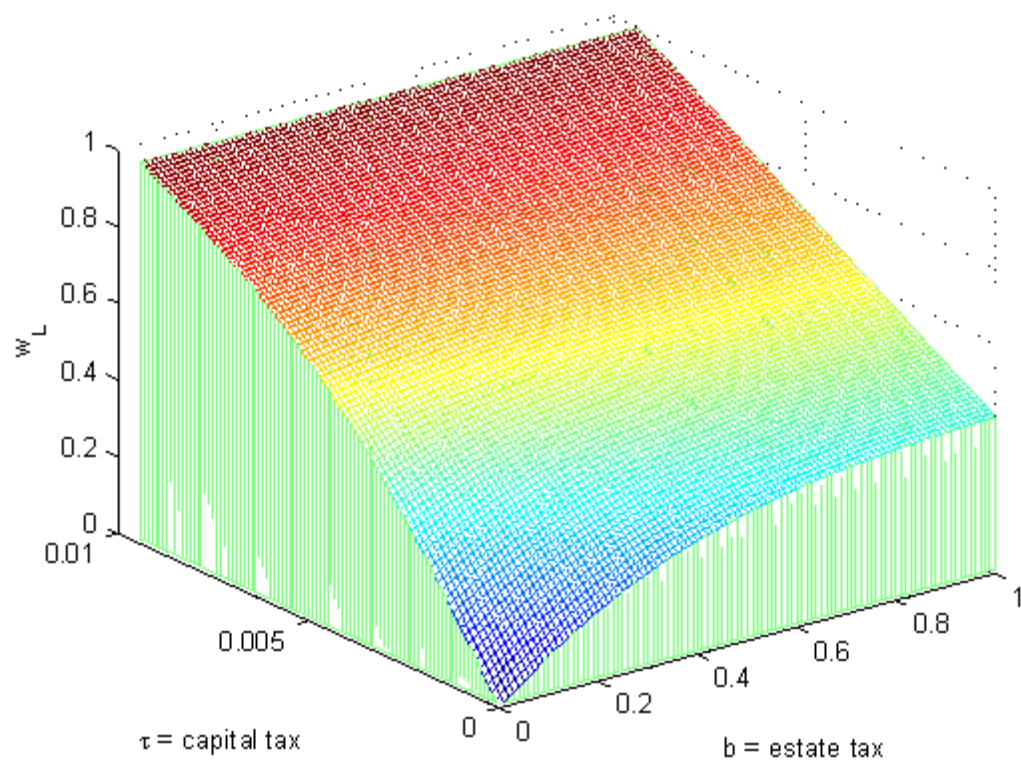
transfers to the young consitute the channel through which the government redistributes wealth in our economy. If we identify such wealth transfers with the discounted value of public education expenditures, for the U.S. in 2003 we obtain 5.9% of GDP.[36] Such expenditures are financed with not only capital and estate tax proceeds but also labor, payroll, indirect and other taxes. If all or a fraction of estate tax collections, which are relatively insignificant, are allocated to finance public expenditure revenues, there remains a share corresponding to $5.9 - 5.5\%$ of GDP to be financed by other taxes. If we assume that public education expenditures are financed by capital taxes in proportion to the share of capital taxes in government tax collections, about 20% (see Auerbach et al (1995), aggregating male and female cohorts from tables 3a-b), then the share of public education expenditures financed by capital and estate taxes are 1.1% to 1.2% of GDP, or about 0.4% of non-human wealth. Accordingly, we set $\tau = 0.004$.

For this parametrization of fiscal policy, with $b = 0.19$ and $\tau = 0.004$, at the stationary distribution of wealth of our calibrated economy, the Pareto exponent is $P = 2.6448$, implying a Gini coefficient of $G = 0.2321$ and a ratio between minimum and average wealth, $\frac{w}{M} = 0.6219$. The US wealth Gini coefficient in 1992 is around 0.78 (from Survey of Consumer Finances data, in Castaneda et al. 2003), in which suggests that the wealth inequality induced by intergenerational transmission can account for almost a third of the observed wealth inequality.[37]

Estate tax flows implied by thecalibration are consistent with thedata. According to OMB Watch (2001), estate taxes in 1996 were about 0.3% of GDP; see also Gale and Slemrod (2000). Also, received bequests, not including inter-vivos transfers, accounted for about 2% of GDP (adjusted to exclude earning), according to Hendricks (2002) in Survey of Consumer Finance data 1989. Our calibration yields pre-tax estates as a fraction wealth at $(1 - b)q(1 - \mu)$, or 1.9% of GDP.

[36] See international tables from UNICEF which report public education expenditures for 2003, at http://stats.uis.unesco.org/TableViewer/tableView.aspx?ReportId=219

[37] The fraction of wealth held in non-annuitized assets that we adopted in the benchmark calibration might be imprecisely measured. To do a sensitivity analysis we compute the parameters of the wealth distribution in the case $\frac{q}{p}(1 - \mu) = 0.2$ and $\frac{q}{p}(1 - \mu) = 0.6$ obtaining, respectively, $P = 1.7142$, $G =$

4.2 Normative effects of fiscal policies

Instead of focusing on inequality, we may take social welfare to be the main target of fiscal policy. This of course requires the choice of a social welfare function.[38]

In the context of an additively separable (utilitarian) welfare criterion, we can inquire into the welfare properties of the stationary distribution of wealth $f(z)$.[39] We can in fact express the social welfare of the agents alive at an arbitrary time t as a function of the Pareto exponent P.

Consider a representative agent who solves the maximization problem (1-2). Her optimal consumption-savings choice path is characterized in Section 2. Given an arbitrary discounted wealth z at time t, her time t discounted utility along the optimal path can be written as (see the derivation in Appendix A):

$$U(z) = \frac{1}{\theta + p} \left(\frac{g\,(1 + p\chi)}{\theta + p} + \ln \eta + p\chi \ln (\eta\chi)\,(1 - b) \right) + \frac{1 + p\chi}{\theta + p} \ln z \qquad (29)$$

It is independent of t. Recall that a fraction $\frac{p-q}{p}$ of the agents have no preferences for

---

$0.4118, w_L = 0.4166$ and $P = 5.9208$, $G = 0.0922, w_L = 0.8311$. For the benchmark calibration of $\frac{q}{p}(1 - \mu) = 0.4$, but with alternative capital taxes of $\tau = .0025$, $\tau = 0.003$ and $\tau = 0.006$ the Gini coefficients in the calibration are, respectively, $G = 0.4525$, $G = 0.4048$, $G = 0.1248$. See also Appendix F.

[38] A large literature has explored the properties of social welfare functions, in particular those that are additively separable in individual utilities and that are increasing in the mean of the distribution of income and decreasing in a measure of its dispersion for all possible income or wealth distributions; see Samuelson (1965) for an early contribution to the subject. Atkinson (1970) and Newbery (1970) demonstrated that if individual utilities are strictly concave there exists no additively separable social welfare function that satisfies these properties; and later Sheshinski (1972) demonstrated that a Rawlsian welfare criterion would indeed satisfy them.

[39] Chipman (1974), restricting his attention to Pareto distributions, showed that with additively separable social welfare functions, increasing the Pareto coefficient (and thus decreasing the Gini coefficient) does indeed increase social welfare if the mean (rather than the lower bound) of the distribution is kept constant. These results however are derived in a static context and cannot be applied directly to our model.

bequests, that is, they have $\chi = 0$. For these agents, given an arbitrary discounted wealth z at time t, their time t discounted utility along the optimal path can be written as:

$$U_0(z) = \frac{1}{\theta + p} \left(\frac{g}{\theta + p} + \ln(p + \theta) \right) + \frac{1}{\theta + p} \ln z$$

The utilitarian social welfare of the agents alive at an arbitrary time, at the stationary wealth distribution $f(z)$ defined by (24), a Pareto distribution with mean M and exponent P, is:

$$
\begin{align}
\Omega(M, \tau, b) &= \frac{q}{p} \int_{\underline{w}}^{\infty} U(z) f(z) dz + \frac{p - q}{p} \int_{\underline{w}}^{\infty} U_0(z) f(z) dz, \tag{30} \\
\text{where } \underline{w} &= \frac{(\tau + bq(1 - \mu)) M}{p - q(1 - \mu)(1 - b) \left(a^* + \frac{P}{P-1}(1 - a^*) \right)} \tag{31} \\
P &= \frac{p - a^* q (1 - \mu)(1 - b)}{p - q(1 - \mu) - \tau} \text{ with } a^* \text{solving (23)} \tag{32}
\end{align}
$$

We can now consider the welfare effects of different fiscal policies, that is, of different combinations of estate taxes b and capital income taxes τ which satisfy government budget balance. A policy (b, τ) affects on the Pareto exponent P of the stationary distribution $f(z)$ as P depends on τ and b. Note that, in a static framework without growth and without a bequest motive, the utilities of agents and the social welfare function do not directly depend on b or on τ except through the Pareto coefficient. Maximizing social welfare would then be equivalent to maximizing P and, given the egalitarian social welfare function, not surprisingly, welfare would be maximized under complete equality: $P = \infty$ and $G = 0$. However this is no longer the case in our dynamic context because both τ and b enter the social welfare function through g and through the bequest motive, in addition to entering through the Pareto coefficient.

The derivatives of the social welfare function with respect to τ and b are reported in the Appendix A. From Proposition 5 we know that when the Pareto exponent is maximized at $\tau = p - q(1 - \mu)$, we have $\frac{\partial P}{\partial b} = 0$. Consequently, for $\chi > 0$ social welfare

would decline in b due to the bequest motive and the optimal b would be zero. If however τ has an interior solution so that $\frac{\partial P}{\partial b} > 0$, we cannot determine whether or not b will be interior. In fact it can be shown by inspecting the derivatives in the Appendix that the value of τ that maximizes social welfare has to be less than $p - q(1 - \mu)$.

To better illustrate the welfare effects of taxes τ and b we can revert to the calibrated economy. To compute social welfare, however, we need to distinctly set q and χ. If we set the proportion of agents who leave bequests, $\frac{q}{p}$ to 0.7,[40] we obtain $q = .0112$ and $1 - \mu = 0.58$, yielding $q(1 - \mu) = 0.0065$. Using $1 - \mu = \chi \frac{p+\theta}{p\chi+1}$, from Proposition 1, we then obtain $\chi = 12.5$. Finally, we set the share of government expenditures out of wealth to $\gamma = 0.1$, corresponding to about 30% of GDP.

Figure 3 shows the plot of the social welfare function as a function of b and τ.

For our benchmark calibration, maximizing social welfare requires setting $b = 0$. The reason is that the negative effect of b on social welfare through its reduction of bequests, given by $-p\chi(1-b)^{-1}$, dominates the positive effect of b on social welfare through the Pareto exponent.[41] Welfare is maximized at $(\tau, b) = (0.0092, 0)$ where the maximum

---

[40] In the Survey of Consumer Finances 1989 only 30% of subjects declare having received an inheritance; see Hendricks (2001). This appears to be a severe under-estimation, however, since *i)* agents seem generally to under-report gifts received (e.g., in the sample of decedents from the HRS/AHEAD survey, studied by Hurd-Smith, 1999, 70% leave a positive estate) and since *ii)* it does not contain inter-vivos transfers. See also Gale and Scholtz (1994).

[41] In the altruistic specification of preferences for bequests studied in Appendix A, this would tend to reduce redistributive taxes at an optimum, as welfare subsidies would induce agents to reduce the investment in assets in the early wealth accumulation stages.

Note also that our normative fiscal policy analysis changes if we restrict to a formulation of "joy of giving" preferences for bequests which depends on gross rather than net bequests. Under logarithmic utility the share of consumption η and the portfolio allocation determined by μ are independent of b, and only the social welfare function is affected. Now however the welfare maximizing $b = 0.990$, is still less than the maximum of 1 but much higher than zero. When agents derive utility from gross bequests, it becomes optimal, given the egalitarian social welfare function, to redistribute wealth with both high capital and high estate taxes.
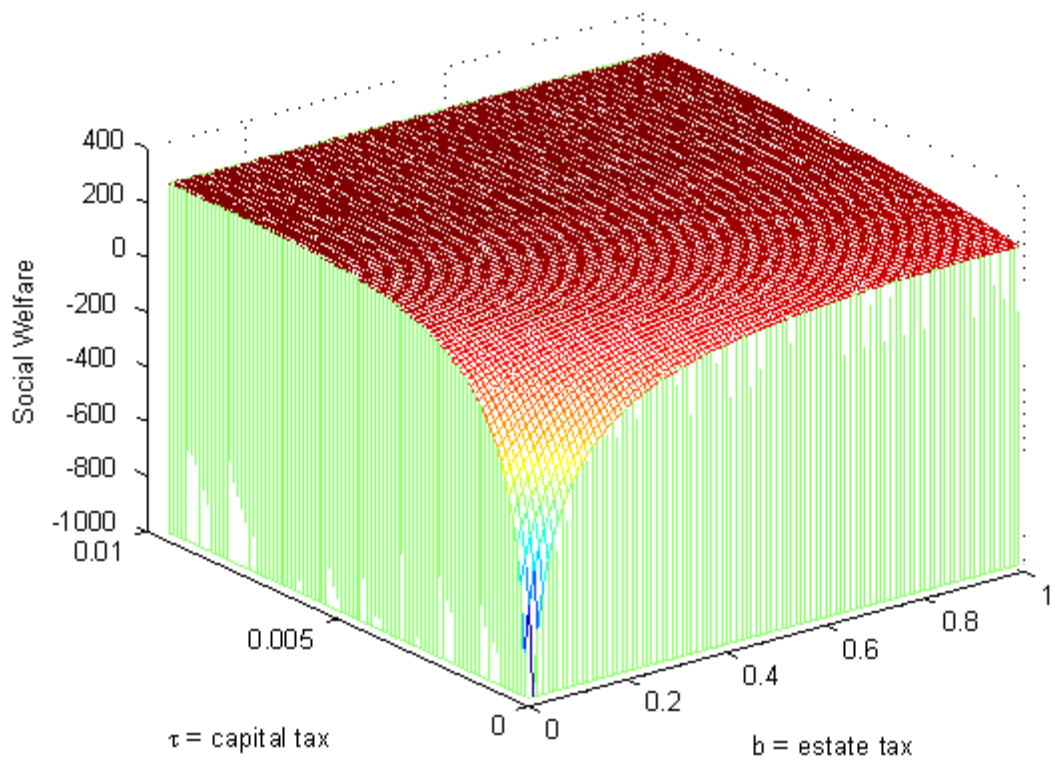
value of the net tax rate τ is $p - q(1 - \mu) = 0.0094$, so τ is indeed interior.

Optimal estate taxes remain at zero, a result robust to varying γ, and capital taxes are interior but close to their maximum allowed value of $p - q(1 - \mu)$, which assures almost complete equality. At the optimum capital tax almost all the population is concentrated just below the mean wealth of 1.[42] The egalitarianism implicit in the social welfare function is implemented through capital taxes rather than through estate taxes. In our calibration however this comes at the expense of growth in the wealth of the agents of 1%. This growth effects is the reason why complete re-distribution $(G = 0)$ is not optimal in our economy.

Not surprisingly, the utilitarian social welfare criterion favors a major wealth redistribution in our economy. In an economy in which e.g., hidden effort or unobservable skills and endogenous labor supply affect individual earnings, as in Atkeson-Lucas (1995), Phelan (1998), or Kocherlakota (2005), unequal wealth distributions naturally arise from optimal tax exercises, as inequality positively affects incentives. It is notable, that even in our context, with no earnings and hence no role for labor supply incentives, optimal taxes do not imply complete redistribution. Furthermore, our social welfare function weighs the well-being of future generations only through the preferences of those currently alive. An alternative approach, due to Caplin and Leahy (2004), is to give weight to future generations in addition to their implicit valuation through their current ancestors. Such a welfare function would put more weight on growth than a standard welfare function and moderate the capital income taxes that impede growth in our model.

5 Discussion and Conclusions

The distribution of wealth in our economy is determined by the intergenerational transmission of wealth. It is skewed, in fact it is Pareto, even without the help of a skewed

---

[42]These properties of optimal taxes are robust to the sensitivity analysis we have performed, with $\frac{q}{p} = .5, .6,$ and $.8$ and with $\chi = 8, 10, 12,$ and 15; see Appendix F for details.

distribution of earnings. A Pareto distribution results endogenously from the interplay and the balancing of two opposite forces, wealth accumulation and redistribution.

The economy we studied is special in several dimensions. First of all, preferences are logarithmic to facilitate the derivation of results in closed form. More importantly, the demographic structure of the economy is characterized by a constant probability of death (independent of age) and an overlapping generations structure (a child is born when a parent dies). While obviously overly restrictive and motivated by the necessity of a stationary demographic structure to solve for consumption-savings paths in closed form, we do not believe that these assumptions invalidate our analysis. Similar results could be obtained in simulations provided agents save enough in old age (which can be obtained tampering with preferences for bequests as in some of the literature surveyed by De Nardi and Cagetti (2005). Also, we have studied a specific class of fiscal policy, consisting of capital income and estate taxes redistributed via means-tested subsidies. As shown in Appendix E, most of our analysis can be extended to the case of lump-sum subsidies, that is, to the case in which subsidies are obtained by all agents in the economy in equal measure, an obviously much less re-distributive welfare policy. It would also be interesting to study consumption taxes. Consumption taxes, differently from capital income taxes, have in fact no growth effects, and hence would allow re-distribution with minimal effects on the growth rate of discounted wealth.

While our economy is special in several dimensions, the inter-generational wealth transmission mechanism studied in our economy operates more generally to generate Pareto distributions of wealth. Consider a dynastic infinite horizon economy in which all agents face a Poisson probability p that their wealth is wiped out unless invested in a protected insurance account. Let $\omega(s, t)$ denote the amount of wealth deposited at time t by an agent born at time s in the protected insurance account. Let r denote the interest rate on wealth at time t (no annuity component). Assume that the insurance account pays a return $r - \delta$, where $\delta - p \geq 0$ is a measure of market imperfection.

The maximization problem of an agent born at time s, in recursive form, is:

$$V\left(w(s,t)\right) = max_{c,\omega} \int_t^\infty e^{(\theta+p)(t-v)} \left(\ln c\left(s,v\right) + pV\left(\omega(s,v)\right)\right) dv$$

subject to

$$\frac{dw\left(s,t\right)}{dt} = rw\left(s,t\right) - \delta\omega - c\left(s,t\right)$$

and the Transversality condition. Proceeding as in the Overlapping Generations economy with altruistic agents (see Appendix A), we can show that

$$c = \theta w, \quad \omega = \frac{p}{\delta}w$$

When $p = \delta$ and insurance is without friction, then all of the wealth is deposited in the insurance account.

It follows immediately that our whole analysis of the dynamics of the distribution of wealth can be extended to this dynastic economy, once the parameter χ is taken to be endogenous and $1 - \mu$ is appropriately redefined.

This dynastic economy might more closely represent the mechanism Pareto called *the rise and fall of elites*, to which he attributed the skewed distribution of wealth he had observed across countries and time series.

References

H.J. Aaron and A.H. Munnell (1992): "Reassessing the Role of Wealth Transfer Taxes," *National Tax Journal*, 45, 119-43.

A. Atkeson and R.E. Lucas (1995): "Efficiency and Equality in a Simple Model of Unemployment Insurance", *Journal of Economic Theory*, 66(1), 64-88.

A.B. Atkinson (1970): "On the Measurement of Inequality," *Journal of Economic Theory*, 2, 244-63.

A.B. Atkinson (2001): 'Top Incomes in the United Kingdom over the Twentieth Century,' mimeo, Nuffield College, Oxford.

A.J. Auerbach, G. Jagadeesh Gokhale, L.J. Kotlikoff; J. Sabelhaus, and D. Weil (1995): "The Annuitization of Americans Resources – A Cohort Analysis", NBER W.P. 5089.

G. S. Becker (1973): "A Theory of Marriage: Part I." *Journal of Political Economy*, 81, 4, 813-846.

G. S. Becker and N. Tomes (1979): "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility," *Journal of Political Economy*, 87, 6, 1153-1189.

K. Bjerke (1970): 'Income and Wage Distributions-Part I: A Survey of the Literature,' *Review of Income and Wealth*, 16(3), 235-52.

O.J. Blanchard (1985): 'Debt, Deficits, and Finite Horizons,' *Journal of Political Economy*, 93, 223-47.

F.D. Blau and L.M. Kahn (2005): 'Do Cognitive Test Scores Explain Higher U.S. Wage Inequality?,' *The Review of Economics and Statistics*, 87, 1, 184-93.

M. Brown, C.C. Channing, and S.H. Chiang (2006a): 'Limiting Densities for a Class of Markov Processes,' mimeo, York University, Toronto.

M. Brown, C.C. Channing, and S.H. Chiang (2006b): 'Wealth and Occupations in a Natural Aristocratic Republic,' mimeo, York University, Toronto.

Burris, Val (2000): "The Myth of Old Money Liberalism: The Politics of the "Forbes" 400 Richest Americans," *Social Problems,* 47, 360-378.

M. Cagetti and M. De Nardi (2000): 'Entrepreneurship, Bequests and the Distribution of Wealth,' mimeo, University of Chicago.

M. Cagetti and M. De Nardi (2003): 'Entrepreneurship, Frictions, and Wealth,' mimeo, University of Chicago.

M. Cagetti and M. De Nardi (2005): 'Wealth Inequality: Data and Models,' Federal Reserve Bank of Chicago, W.P. 2005-10.

F.P. Cantelli (1921): 'Sulla Deduzione delle Leggi di Frequenza da Considerazioni di Probabilita',' *Metron*, 83-91.

A. Caplin and J. Leahy (2004): "The Social Discount Rate", *Journal of Political Economy,* 112, 2004, 1257-1268.

A. Castaneda, J. Diaz-Gimenez, and J.V. Rios-Rull (2003): 'Accounting for the U.S. Earnings and Wealth Inequality,' *Journal of Political Economy*, 111, 4, 818-57.

D.G. Champernowne (1953): 'A Model of Income Distribution,' *Economic Journal*, 63, 318-51.

J.S. Chipman (1974): 'The Welfare Ranking of Pareto Distributions,' *Journal of Economic Theory,* 9, 275-82.

J.S. Chipman (1976): *The Paretian Heritage,* Revue europeenne des sciences sociales et Cahiers Vilfredo Pareto, Vol. 14, No. 37, pp. 65-171,

    http://www.econ.umn.edu/~jchipman/HERITAGE.pdf .

F. Clementi and M. Gallegati (2004): 'Power Law Tails in the Italian Personal Income Distribution,' forthcoming, *Physica A.*

J.K. Dagsvik and B.H. Vatne (1999): 'Is the Distribution of Income Compatible with a Stable Distribution?'

A. Das and S. Yarlagadda (2003): 'A Distribution Function Analysis of Wealth Distribution, arXiv:con-mat/0310343.

M. De Nardi (2004): 'Wealth Inequality and Intergenerational Links,' *Review of Economic Studies*, 71, 743-768.

A. Deaton, P.O. Gourinchas, and C. Paxson (2000): 'Social Security and Inequality over the Life Cycle,' NBER Working Paper 7570.

Elwood, Paul, S. M. Miller, Marc Bayard, Tara Watson, Charles Collins, Chris Hartman, (1997): *Born on Third Base: The sources of wealth of the 1996 Forbes 400*, Boston: Unified for a Fair Economy. For an executive summary see

    http://www.faireconomy.org/press/archive/Pre_1999/forbes_400_study.html .

S.J. Farlow (1982): *Partial Differential Equations for Scientists and Engineers,* New York, Dover.

D. Feenberg and J. Poterba (2000): 'The Income and Tax Share of Very High Income Household: 1960-1995,' *American Economic Review*, 90, 264-70.

E. Fermi (1949): 'On the Origin of Cosmic Radiation,' *Physics Review*, 75, 1169-74. J. Friedman and R. Carlitz (2005): 'Estate Tax Reform Could Raise Much-Needed Revenue, Center on Budget and Policy Priorities

J. Friedman and R. Carlitz (2005): 'Estate Tax Reform Could Raise Much-Needed Revenue, Center on Budget and Policy Priorities, at Budget), available at

http://www.ombwatch.org/article/articleview/400/1/116?TopicID=1

A. Fujihara, T. Ohtsuki, and H. Yamamoto (2004): 'Power-Law Tails in Nonstationary Stochastic Processes with Asymmetrically Multiplicative Interactions,' *Physical Review E*, 70, 031106.

X. Gabaix, (1999): 'Zipf's Law for Cities: An Explanation,' *Quarterly Journal of Economics*, 113, 3, 739-767.

X. Gabaix, P. Gopikrishnan, V. Plerou, and H.E. Stanley (2003): 'Theory of Power Law Distributions in Financial Market Fluctuations' *Nature*, 423, 267-70.

W. G. Gale and J. K. Scholtz (1994): "Intergenerational Transfers and the Accumulation of Wealth,' *Journal of Economic Perspectives*, 8, 145-160.

I.I. Gikhman and A.V. Skorokhod (1969): *Introduction to the Theory of Random Processes,* Philadelphia, W.B. Saunders.

J. Diaz Gimenez, V. Quadrini, and J.V. Rios Rull (1997): '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.

J. Heckman and E. Vytlacil (2001): 'Identifying the Role of Cognitive Ability in Explaining the Level of and Chance in the Return to Schooling,' *The Review of Economics and Statistics*, 83, 1, 1-12.

L. Hendricks (2001): "Bequests and Retirement Wealth in the United States," Arizona State University, http://www.lhendricks.org/Research/bequdata_paper.pdf

R. Herrnstein and C. Murray (1994): *The Bell Curve,* New York, Free Press.

A.W. Hood (2003): *Partial Differential Equations (10 Lectures)*, School of Mathematical and Computational Sciences, University of St. Andrews, mimeo, http://www-solar.mcs.st-and.ac.uk/~alan/MT2003/PDE/PDE.html.

M. Huggett (1996): *Journal of Monetary Economics,* "Wealth Distribution in Life-Cycle Economies," 38, 469-494.

M. Hurd and J.P. Smith (1999): "Savings of the Elderly and Desired Bequests," mimeo, Rand Corporation.

M. Kalecki (1945): "On the Gibrat Distribution," *Econometrica,* 13, 161-170.

K. Kawamura and N. Hatano (2003): 'Models of Universal Power-Law Distributions,' arXiv:con-mat/0303331.

O. S. Klassa, O. Bihama, M. Levy, O. Malcaia and S. Solomon (2006), "The Forbes 400 and the Pareto Wealth Distribution," *Economics Letters,* 90(2), 290-295

N. Kocherlakota (2005): "Zero Expected Wealth Taxes: A Mirrlees Approach to Dynamic Optimal Taxation," *Econometrica*, 73, 1587-1621.

L. J. Kotlikoff and L. H. Summers (1981) "The Role of Intergenerational Transfers in Aggregate Capital Accumulation," *Journal of Political Economy*, 89, 706-732.

P. Krusell and A. A. Smith (1998): 'Income and Wealth Heterogeneity in the Macroeconomy,' *Journal of Political Economy*, 106, 867-896.

S. Kuznets (1955): 'Economic Growth and Economic Inequality,' *American Economic Review*, 45, 1-28.

J. Laitner (2001); "Secular Change in wealth Inequality and Inheritance," *Economic Journal*, 111, 691-721.

R.J. Lampman (1962): *The Share of Top Wealth-Holders in National Wealth, 1922-1956*, Princeton, NJ, NBER and Princeton University Press.

M. Levy (2003): 'Are Rich People Smarter?' *Journal of Economic Theory*, 110, 42-64.

R. E. Lucas (1978): "On the Size Distribution of Business Firms." *The Bell Journal of Economics,"* 9, 508-523.

E.G.J. Luttmer (2004): 'The Size Distribution of Firms in an Economy with Fixed and Entry Costs,' Federal Reserve Bank of Minneapolis W.P. 633.

E.G.J. Luttmer (2005): 'Growth and the Size Distribution of Firms,' University of Minnesota, mimeo.

O. Malcai, O. Biham, P. Richmond, and S. Solomon (2002): 'Theoretical Analysis and Simulations of the Generalized Lotka-Volterra Model,' *Physical Review E*, 66, 031102.

B. Mandelbrot (1960):"The Pareto-Levy Law and the Distribution of Income." *International Economic Review,* 1, 79-106.

B. Mandelbrot (1962): 'Paretian Distributions and Maximization,' *Quarterly Journal of Economics*, 76(1), 57-85.

J. Mincer (1958): "Investment in Human Capital and Personal Income Distribution," *Journal of Political Economy,* 66, 281-302.

R.N. Mantegna and H.E. Stanley (2000): *An Introduction to Econophysics*, Cambridge, U.K., Cambridge University Press.

C. Moriguchi and E. Saez (2005): 'The Evolution of Income Concentration in Japan, 1885-2002: Evidence from Income Tax Statistics,' mimeo, University of California, Berkeley.

C. Morisson (2000): 'Historical Perspectives on Income Distribution: The Case of Europe,' in A. Atkinson and F. Bourguignon (eds.), *Handbook of Income Distribution*, Amsterdam, North Holland.

D. Newbery (1970): 'A Theorem on the Measurement of Inequality,' *Journal of Economic Theory*, 2, 264-66.

M. Nirei and W. Souma (2004): 'Two Factor Model of Income Distribution Dynamics,' mimeo, Utah State University.

J.P. Nolan 2005): *Stable Distributions; Models for Heavy Tailed Data*, mimeo, Math/Stat Department, American University, Washington, D.C.

OMB Watch, "Estate Tax and Government Revenue," 2001,

  http://www.ombwatch.org/article/articleview/400/1/116?TopicID=1

V. Pareto (1897): *Cours d'Economie Politique*, Vol. II, F. Rouge, Lausanne.

V. Pareto (1901): "Un' Applicazione di teorie sociologiche", *Rivista Italiana di Sociologia.* 5. 402-456, translated as *The Rise and Fall of Elites: An Application of Theoretical Sociology* , Transaction Publishers, New Brunswick, New Jersey, (1991).

V. Pareto (1909): *Manuel d'Economie Politique*, V. Giard et E. Brière, Paris.

C. Phelan (1998): "On the Long-Run Implications of Repeated Moral Hazard," *Journal of Economic Theory*, 79(2), 174-91.

E.S. Phelps and R.A. Pollak (1968): 'On Second-best National Savings and Game-equilibrium Growth,' *Review of Economic Studies*, 35, 185-99.

T. Piketty (2001): 'Income Inequality in France, 1901-1998,' *Journal of Political Economy.*

T. Piketty and E. Saez (2003): 'Income Inequality in the United States, 1913-1998,' *Quarterly Journal of Economics*, CXVIII, 1, 1-39.

A.D. Polyanin and A.V. Manzhirov (1998): *Handbook of Integral Equations,* Boca Raton, Florida, CRC Press.

V. Quadrini (2000): 'Entrepreunership, Savings and Social Mobility,' *Review of Economic Dynamics*, 3, 1-40.

W.J. Reed and M. Jorgensen (2003): 'The Double Pareto-Lognormal Distribution: A New Parametric Model for Size Distributions,' mimeo, University of Victoria.

Roy, A. D. (1950) 'The Distribution of Earnings and Individual Output," *The Ecoomic Journal*,' 60, 489-505.

R.S.G. Rutherford (1955): 'Income Distribution: A New Model,' *Econometrica*, 23, 277-94.

E. Saez and M. Veall (2003): 'The Evolution of Top Incomes in Canada,' NBER Working Paper 9607.

P.A. Samuelson (1965): 'A Fallacy in the Interpretation of the Pareto's Law of Alleged Constancy of Income Distribution,' *Rivista Internazionale di Scienze Economiche e Commerciali*, 12, 246-50.

N. Scafetta, S. Picozzi, and B. West (2002): 'Pareto's Law: A Model of Human Sharing and Creativity,' arXiv:cond-mat/0209373 v1.

E. Sheshinski (1972): 'Relation between a Social Welfare Function and the Gini Index of Income Inequality,' *Journal of Economic Theory*, 4, 98-100.

H.A. Simon (1955): 'On a Class of Skew Distribution Function,' *Biometrika*, 52, 425-40.

H.A. Simon and C.P. Bonini (1958): 'The Size Distribution of Business Firms,' *American Economic Review*, 48, 607-17.

S. Solomon (1999): 'Generalized Lotka-Volterra (GLV) Models,' arXiv:con-mat/9901250 v1.

G. Solon (1999): 'Mobility Within and Between Generations,' mimeo

D. Sornette (2000): *Critical Phenomena in Natural Sciences,* Berlin, Springer Verlag.

W. Souma, Y. Fujiwara, and H. Aoyama (2001): 'Small-World Effects in Wealth Distribution,' arXiv:con-mat/0108482.

J.C. Strikwerda (2004): *Finite Difference Schemes and Partial Differential Equations,* Philadelphia, SIAM.

F.G. Tricomi (1957): *Integral Equations,* New York, Interscience (Dover edition, 1985).

M. Yaari (1965): 'Uncertain Lifetime, Life Insurance, and the Theory of the Consumer,' *Review of Economic Studies*, 32, No. 2, 137-150.

N. Wang (2006): 'An Equilibrium Model of Wealth Distribution,' *Journal of Monetary Economics*, forthcoming.

H.O.A. Wold and P. Whittle (1957), 'A Model Explaining the Pareto Distribution of Wealth,' *Econometrica*, 25, 4, 591-5.

E. Wolff (1987): 'Estimates of Household Wealth Inequality in the U.S., 1962-1983,' *The Review of Income and Wealth*, 33, 231-56.

E. Wolff (2004): 'Changes in Household Wealth in the 1980s and 1990s in the U.S.,' mimeo, NYU.

Appendix A: Proofs - for completeness

Proof of Prop. 1. The dynamic equation for wealth accumulation is

$$\frac{dw(s,t)}{dt} = (r+p-\tau)w(s,v) - p\omega(s,v) - c(s,v)$$

First order conditions include

$$\omega(s,t) = \chi c(s,t) \tag{33}$$

$$\dot{c}(s,t) = (r-\tau-\theta)c(s,t) \tag{34}$$

The aggregate dynamics for the agent can then be written as:

$$\dot{w}(t,s) = (r+p-\tau)w(t,s) - (p\chi+1)c(s,v)$$

Postulating $c = \eta w$, after some algebra,

$$\frac{dw(s,t)}{dt} = ((r+p-\tau) - \eta(p\chi+1))w(s,v) \tag{35}$$

So that, equating (34) and (35) we verify that in fact

$$c = \eta w, \quad \text{with } \eta = \frac{(p+\theta)}{p\chi+1} \tag{36}$$

Furthermore, by (33),

$$\omega = \chi\eta w,$$

Finally, using

$$\dot{w}(t,s) = (r+p-\tau)w(t,s) - (p\chi+1)\eta w(t,s)$$

and

$$\eta = \frac{(p+\theta)}{p\chi+1}$$

we can solve for the growth of the agent's wealth, which we denote g:

$$g = r - \tau - \theta \tag{37}$$

Optimal Consumption-Savings for Altruistic Agents. The maximization problem of an agent born at time s, in recursive form, is:

$$V\left(w(s,t)\right) = max_{c,\omega} \int_t^\infty e^{(\theta+p)(t-v)} \left(\ln c\left(s,v\right) + p\alpha V\left(\omega(s,v)\right)\right) dv$$

subject to

$$\frac{dw\left(s,t\right)}{dt} = (r+p)\, w\left(s,t\right) - p\omega - c\left(s,t\right)$$

and the Transversality condition

We guess

$$V\left(\omega\right) = A + B\ln\omega$$

Under the guess the Hamiltonian is

$$H = \ln c + pA\alpha + pB\alpha \ln\omega + \lambda\left((r+p)w - p\omega - c\right)$$

and the associated First Order Conditions are

$$
\begin{aligned}
c^{-1} &= \lambda;\ \frac{\dot{c}}{c} = -\frac{\dot{\lambda}}{\lambda} \\
\frac{\alpha B}{\omega} &= \lambda \\
\dot{\lambda} &= \lambda\left(-r-p+p+\theta\right) = -\lambda\left(r-\theta\right) \\
\dot{c} &= c\left(r-\theta\right)
\end{aligned}
$$

with the wealth accumulation constraint.

Under the guess we can show that

$$c = \eta w, \quad \omega = \alpha B \eta w$$

with

$$\eta = \frac{p + \theta}{p \alpha B + 1}$$

Also, under the guess, we can compute the value $V(z)$ as:

$$V(z) = \text{ constant } + \frac{1 + p \alpha B}{\theta + p} \ln z$$

(see the "Derivation of the discounted utility along the optimal path" above). The guess is verified therefore for[43]

$$B = \frac{1 + p \alpha B}{\theta + p}$$

that is

$$B = \frac{1}{\theta + p(1 - \alpha)}$$

We conclude that:

$$\eta = \frac{p + \theta}{p \alpha \frac{1}{\theta + p(1 - \alpha)} + 1} = \theta + p(1 - \alpha)$$

and so

$$c = (\theta + p(1 - \alpha)) \, w, \quad \omega = \alpha w$$

As noted in the text (see especially footnote 22, when parents are purely altruistic, they optimally account for welfare transfers when choosing bequests. In this case, the consumption-savings plan obtained in Proposition 1 is not optimal.

It is straightforward to show, following the analysis just above, that this case is homeomorphic to one in which agents have non-convex preferences for bequest of the following kind:

---

[43]Naturally we also need to require constant$= A$.

$$\phi\left(\omega\left(s,v\right)\right) = \begin{cases} \chi \ln \underline{w}(v) & for\ 0 \le (1-b)\,\omega\left(s,v\right) \le \underline{w}(v) \\ \chi \ln(1-b)\omega\left(s,v\right) & for\ \underline{w}(v) \le (1-b)\omega\left(s,v\right) \end{cases} \tag{38}$$

The optimal consumption-saving path (no capital tax for simplicity) in this case has the following properties:

1. there exists a finite wealth level $\hat{w}$ such that the consumption-savings path in Proposition 1 is optimal for $w > \hat{w}$; moreover,

$$\hat{w} = \left(\frac{w(\hat{t})}{(1-b)\chi\eta}\right)$$

$$\underline{w}(\hat{t}) = \underline{w}(0)e^{g'\hat{t}} =$$

$$e^{(r+p)\hat{t}}w\left(0\right)\left(1 - \frac{\frac{\eta}{p+\theta}\left(e^{(p+\theta)\hat{t}} - 1\right)}{1 + \frac{\eta}{p+\theta}\left(e^{(p+\theta)\hat{t}} - 1\right)}\right)$$

implying that

$$\hat{w} \to \underline{w} \text{ as } \underline{w} \to 0$$

and hence that the consumption-savings path in Proposition 1 is an approximation to the optimal path for $\underline{w}$ small enough.

2. At $\hat{w}$ bequests ω jump from zero to positive, consumption c is smooth, wealth is continuous but has a kink, as $\dot{w}$ jumps at $\hat{w}$.

3. The growth rate of individual wealth is not constant; let it be denoted $g(w)$:

$$g(w) = \begin{cases} r+p-\eta\left(\frac{(1+p\chi)}{(p\chi)e^{(p+\theta)\hat{t}}+1}\right) & \text{at } w = \underline{w} \\ r+p-\eta & \text{at the left limit of } \hat{w} \\ r+p-\eta\left(1+p\chi\right) & \text{at the right limit of } \hat{w} \end{cases}$$

and it is decreasing in wealth w.

Derivation of the PDE, equation (9). Consider the Chapman-Kolmogorov equation which governs the dynamics of $f(w,t)$. Let $w_1 > w_l(t)$. The mass of wealth in the interval (w_1, w) at time $t + \Delta$ is $\int_{w_1}^{w} f(w, t + \Delta)\, dw$. At a first order approximation this mass has two components. First, since individual wealth grows at rate g, it contains the mass of agents who have wealth in the interval $((1 - g\Delta)\, w_1, (1 - g\Delta)\, w)$ at time t and are alive at $t + \Delta$. Secondly, through the boundary condition it contains the contribution of those newborns who inherit a fraction of their parents' wealth: the newborns at time t who do not inherit from their parents, or whose inheritance fall below $w_l(t)$ add to the density at $w_l(t)$).

Summarizing, the Chapman-Kolmogorov equation can then formally be written as:

$$\int_{w_1}^{w} f(w, t + \Delta)\, dw = (1 - p\Delta) \int_{(1-g\Delta)w_1}^{(1-g\Delta)w} f(w, t)\, dw + q\Delta \int_{\sigma(w_1)}^{\sigma(w)} f(w, t)\, dw \quad + o(\Delta)$$

Differentiating with respect to w and ignoring second-order terms (terms in Δ^2),

$$f(w, t + \Delta) = (1 - p\Delta)(1 - g\Delta) f((1 - g\Delta w), t) + q\Delta \frac{\partial \sigma(w)}{\partial w} f(w\sigma(w), t))$$

Rearranging,

$$\frac{f(w, t + \Delta) - f(w, t)}{\Delta} =$$

$$= \frac{f((1 - g\Delta w), t) - f(w, t) - (\Delta p + \Delta g) f((1 - g\Delta w), t) + q\Delta \frac{\partial \sigma(w)}{\partial w} f(w\sigma(w), t))}{\Delta}$$

and, letting $\Delta \to 0$,

$$\frac{\partial f(w, t)}{\partial t} = -(p + g) f(w), t) + q \frac{\partial \sigma(w)}{\partial w} f(w\sigma(w), t)) - g \frac{\partial f(w, t)}{\partial w}$$

Derivation of the boundary condition, (11). The two terms of (11) are, respectively, the density of the newborns with no inheritance and the density of the newborn with inheritance lower that $\underline{w}$.

The first term of (11) can be derived from the age distribution. In particular the density of newborn agents (agents of age $a = 0$) with no inheritance is $p - q$. The wealth $w(a)$ of an agent of age a born with wealth $\underline{w}$ is $w(a) = \underline{w}e^{ga}$. Operating the appropriate change of variable to obtain the distribution of wealth from the distribution of age, and evaluating at $w = \underline{w}$, we obtain $\frac{p-q}{g}\frac{1}{\underline{w}}$. The second term is straightforwardly derived.

Proof of Lemma 1. To solve (9) under (14) and (13) we apply the "method of characteristics" as detailed in Appendix C. Let the characteristic space (τ, t) be defined by

$$\frac{dz}{d\tau} = (g - g')z, \quad \frac{dt}{d\tau} = 1.$$

Let $z(0) = m$ and $t(0) = 0$. In the characteristic space the PDE (9) is then reduced to the following differential equation:

$$\frac{d\left(f\left(z(\tau), \tau\right)\right)}{d\tau} = -\left(p + g - g'\right) f\left(z(\tau), \tau\right) + q\frac{\partial \sigma(z)}{\partial z} f\left(\sigma(z(\tau)), \tau\right) \tag{39}$$

It can be verified that (39) has solution:

$$f\left(z\left(\tau\right), \tau\right) = e^{-(p+g-g')\tau} f\left(m, 0\right) + \int_0^\tau q\frac{\partial \sigma(z)}{\partial z} f\left(\sigma(z(\eta)), \eta\right) e^{(p+g-g')(\eta-\tau)} d\eta \tag{40}$$

The characteristic space is split along the characteristic $z = \underline{w}e^{(g-g')\tau}$. In particular, for $z \geq \underline{w}e^{(g-g')\tau}$ the solution to the PDE is determined by the initial condition, while for $z < \underline{w}e^{(g-g')\tau}$ the solution is instead determined by the boundary condition through the inverse transformation $\tau(z, y) = \ln\frac{z}{y}\frac{1}{(g-g')}$. Then, substituting back into the original

space (z,t), we obtain

$$f(z,t) =$$

$$
\begin{cases}
\left(\frac{z}{\underline{w}}\right)^{-\frac{p}{g-g'}-1} f\left(\underline{w}, t - \tau\left(z, \underline{w}\right)\right) + \\
\quad + q \int_{\underline{w}}^{z} \frac{\partial \sigma(y)}{\partial y} f\left(\sigma(y), t - \tau\left(z, y\right)\right)(y)^{\left(\frac{p}{g-g'}\right)}(g-g')^{-1}(z)^{-\left(\frac{p}{g-g'}+1\right)} dy & \text{for } z \in \left(\underline{w}, \underline{w}e^{(g-g')t}\right) \\[2em]
e^{-(p+g-g')t} h\left(ze^{(g-g')t}\right) \\
\quad + q \int_{\underline{w}}^{z} \frac{\partial \sigma(y)}{\partial y} f\left(\sigma(y), t - \tau\left(z, y\right)\right)(y)^{\left(\frac{p}{g-g'}\right)}(g-g')^{-1}(z)^{-\left(\frac{p}{g-g'}+1\right)} dy & \text{for } z \geq \underline{w}e^{(g-g')t}
\end{cases}
$$

Proof of Lemma 2. Consider the dynamics of $f(z,t)$ as characterized by (15) in Lemma 1. Consider discounted wealth levels $z \geq \underline{w}e^{(g-g')t}$. In this region, the density an any time t is

$$e^{-(p+g-g')t} h\left(ze^{(g-g')t}\right) + q \int_{\underline{w}}^{z} \frac{\partial \sigma(y)}{\partial y} f\left(\sigma(y), t - \tau\left(z, y\right)\right)(y)^{\left(\frac{p}{g-g'}\right)}(g-g')^{-1}(z)^{-\left(\frac{p}{g-g'}+1\right)} dy$$

Notice that, if $\sigma(y) = y$, the density in the region $z \geq \underline{w}e^{(g-g')t}$ at time t is larger than in the case $\sigma(y) > y$. But, when $\sigma(y) = y$ (15) can be easily solved to obtain that

$$f(z,t) = e^{-(p-q+g-g')t} h\left(ze^{-(g-g')t}\right), \quad \text{for } z \geq \underline{w}e^{(g-g')t}.$$

It is now straightforward to notice that $= e^{-(p-q+g-g')t} h\left(ze^{-(g-g')t}\right)$ vanishes for $t \to \infty$.

We conclude that at the stationary distribution the whole mass is in the region $z \in \left(\underline{w}, \underline{w}e^{(g-g')t}\right)$. As a consequence, then, from (15),

$$f(z) = \left(\frac{z}{\underline{w}}\right)^{-\frac{p}{g-g'}-1} f\left(\underline{w}\right) + q \int_{\underline{w}}^{z} \frac{\partial \sigma(y)}{\partial y} f\left(\sigma(y)\right)(y)^{\left(\frac{p}{g-g'}\right)}(g-g')^{-1}(z)^{-\left(\frac{p}{g-g'}+1\right)} dy$$

Proof of Prop. 2. Substituting (19), (8), and $\mu = 1$, $q = 0$ into (15) reduces it to

$$
f(z,t)
\begin{cases}
\frac{p}{p-\tau} \frac{1}{\underline{w}} \left(\frac{z}{\underline{w}}\right)^{-\left(\frac{p}{p-\tau}+1\right)} & \text{for } z \in \left(\underline{w}, \underline{w}e^{(p-\tau)t}\right) \\[1em]
e^{-(p+p-\tau)t} h\left(z\right) & \text{for } z \geq \underline{w}e^{(g-g')t}
\end{cases}
\tag{41}
$$

Furthermore, using

$$f(z,t) = e^{-(p+\rho-\tau)t} h\left(ze^{-(\rho-\tau)t}\right), \quad \text{for } z \geq \underline{w}e^{(g-g')t}.$$

it follows that $f(z,t)$ vanishes for $t \to \infty$ and hence that $f(z)$ is ergodic.

Finally, the mean of the stationary distribution $f(z)$ is finite since $\frac{p}{\rho-\tau} > 1$.

Proof of Prop. 3. Consider $z \in \left(\underline{w}, \underline{w}e^{(g-g')t}\right)$. In this range, substituting (19), and $b = 0$, in (15), it follows that $f(z,t)$ is stationary (independent of t), and hence it satisfies the integral equation (16), which in this case takes the form:

$$f(z) = \left(\frac{z}{\underline{w}}\right)^{-\left(\frac{p}{g-g'}+1\right)} f(\underline{w}) + q \int_{\underline{w}}^{z} (y)^{\left(\frac{p}{g-g'}\right)} (g-g')^{-1} (z)^{-\left(\frac{p}{g-g'}+1\right)} f(y)\, dy \qquad (42)$$

This is a Volterra integral equation of the second type, with separable kernel, for which a closed form solution exists and is discussed in Appendix D. Applying this solution, we obtain,

$$\begin{aligned}
f(z) &= \left(\frac{z}{\underline{w}}\right)^{-\left(\frac{p}{g-g'}+1\right)} f(\underline{w}) \\
&\quad + q\,(g-g')^{-1} \int_{\underline{w}}^{z} z^{-\left(\frac{p}{g-g'}+1-q(g-g')^{-1}\right)} j^{\left(\frac{p}{g-g'}-q(g-g')^{-1}\right)} \left(\frac{j}{\underline{w}}\right)^{-\left(\frac{p}{g-g'}+1\right)} f(\underline{w})\, dj
\end{aligned} \qquad (43)$$

Straightforward algebraic manipulations, together with (19), are now enough to produce the closed form solution for the stationary distribution $f(z)$.

Furthermore, as before, using

$$f(z,t) = e^{-(p+\rho-q-\tau)t} h\left(ze^{-(\rho-q-\tau)t}\right), \quad \text{for } z \geq \underline{w}e^{(g-g')t}.$$

it follows that $f(z,t)$ vanishes for $t \to \infty$ and hence that $f(z)$ is ergodic.

Finally, the mean of the stationary distribution $f(z)$ is finite since $\frac{\rho-q}{\rho-q-\tau} > 1$.

Proof of Prop. 4. The integral equation in this case, after the transformation

$j = \sigma(y) = \frac{y}{(1-\mu)(1-b)}$ is reduced to:

$$f(z) = \tag{44}$$

$$\left(\frac{z}{\underline{w}}\right)^{-\left(\frac{p}{g-g'}+1\right)} f(\underline{w}) + q(g-g')^{-1} \cdot$$

$$\cdot \int_{\frac{\underline{w}}{(1-\mu)(1-b)}}^{\frac{z}{(1-\mu)(1-b)}} f(j) \left[((1-\mu)(1-b)j)^{\left(\frac{p}{g-g'}\right)} z^{-\left(\frac{p}{g-g'}+1\right)} \right] dj$$

where $g - g' = p - q(1-\mu) - \tau$. We guess:

$$f(z) = \frac{p - aq(1-\mu)(1-b)}{p - q(1-\mu) - \tau} \, \underline{w}^{\frac{p-aq(1-\mu)(1-b)}{p-q(1-\mu)-\tau}} \, z^{-\left(\frac{p-aq(1-\mu)(1-b)}{p-q(1-\mu)-\tau}+1\right)} \tag{45}$$

and substitute into the integral equation. Let $f(\underline{w}) = \frac{p-aq(1-\mu)(1-b)}{p-q(1-\mu)-\tau} \, \frac{1}{\underline{w}}$. We obtain

$$\frac{z}{\underline{w}}^{-\left(\frac{p-aq(1-\mu)(1-b)}{p-q(1-\mu)-\tau}+1\right)} f(\underline{w}) =$$

$$\frac{z}{\underline{w}}^{-\left(\frac{p}{g-g'}+1\right)} f(\underline{w}) + q(g-g')^{-1} \, (z)^{-\left(\frac{p}{g-g'}+1\right)} \cdot$$

$$\cdot \int_{((1-\mu)(1-b))^{-1}\underline{w}}^{((1-\mu)(1-b))^{-1}z} \left(\frac{j}{\underline{w}}\right)^{-\left(\frac{p}{g-g'}+1-aq(g-g')^{-1}(1-\mu)(1-b)\right)} f(\underline{w}) \left((1-\mu)(1-b)j\right)^{\frac{p}{g-g'}} dj$$

and, after some algebraic manipulations,

$$\frac{z}{\underline{w}}^{-\left(\frac{p-aq(1-\mu)(1-b)}{p-q(1-\mu)-\tau}+1\right)} f(\underline{w}) = \frac{z}{\underline{w}}^{-\left(\frac{p}{g-g'}+1\right)} f(\underline{w}) +$$

$$+ q(g-g')^{-1}((1-\mu)(1-b))^{\frac{p}{g-g'}} \frac{z}{\underline{w}}^{-\left(\frac{p}{g-g'}+1\right)} \underline{w}^{-aq(g-g')^{-1}((1-\mu)(1-b))} f(\underline{w}) \cdot$$

$$\cdot \int_{((1-\mu)(1-b))^{-1}\underline{w}}^{((1-\mu)(1-b))^{-1}z} j^{-1+aq(g-g')^{-1}(1-\mu)(1-b)} dj$$

and hence

$$\frac{z}{\underline{w}}^{-\left(\frac{p-aq(1-\mu)(1-b)}{p-q(1-\mu)-\tau}+1\right)} f(\underline{w}) =$$

$$= \frac{z}{\underline{w}}^{-\left(\frac{p}{g-g'}+1\right)} f\left(\underline{w}\right) \left(\begin{array}{c} 1 + a^{-1}\underline{w}^{-aq(g-g')^{-1}(1-\mu)(1-b)} \left((1-\mu)(1-b)\right)^{\frac{p-aq(1-\mu)(1-b)}{g-g'}-1} \\ \cdot \left(z^{aq(g-g')^{-1}(1-\mu)(1-b)} - \underline{w}^{aq(g-g')^{-1}(1-\mu)(1-b)} \right) \end{array} \right)$$

$$= \frac{z}{\underline{w}}^{-\left(\frac{p}{g-g'}+1\right)} f\left(\underline{w}\right) \left(\begin{array}{c} 1 + a^{-1} \left((1-\mu)(1-b)\right)^{\frac{p-aq((1-\mu)(1-b))}{g-g'}-1} \\ \left(\frac{z}{\underline{w}}^{aq(g-g')^{-1}(1-\mu)(1-b)} - 1 \right) \end{array} \right)$$

Let $a^{-1}\left((1-\mu)(1-b)\right)^{\frac{p-aq(1-\mu)(1-b)}{g-g'}-1} = 1$, or

$$a = \left((1-\mu)(1-b)\right)^{\frac{p-aq(1-\mu)(1-b)}{g-g'}-1}$$

This is fixed point equation which has a unique solution, $a^* < 1$. In fact, it is easily checked that $\left((1-\mu)(1-b)\right)^{\frac{p-aq((1-\mu)(1-b))}{g-g'}-1} = \left((1-\mu)(1-b)\right)^{\frac{p-aq(1-\mu)(1-b)}{p-q((1-\mu))-\tau}-1}$ is strictly positive for $a = 0$, it has a negative derivative with respect to a, and it is less than 1 for $a = 1$. Consequently,

$$\frac{z}{\underline{w}}^{-\left(\frac{p-a^*q((1-\mu)(1-b))}{g-g'}+1\right)} f\left(\underline{w}\right) =$$

$$= \frac{z}{\underline{w}}^{-\left(\frac{p}{g-g'}+1\right)} f\left(\underline{w}\right) \left(1 + \left(\frac{z}{\underline{w}}^{a^*q(g-g')^{-1}(1-\mu)(1-b)} - 1 \right) \right)$$

$$= \frac{z}{\underline{w}}^{-\left(\frac{p-a^*q((1-\mu)(1-b))}{g-g'}+1\right)} f\left(\underline{w}\right)$$

and the guess is verified.

Finally, the mean of the stationary distribution $f(z)$ is finite since $\frac{p-a^*q(1-\mu)(1-b)}{p-*q(1-\mu)-\tau} > 1$.

Since we cannot solve explicitly for the dynamics of $f(z,t)$ in the general case under consideration, we cannot directly show ergodicity of $f(z)$, as in the previous special cases studied in Propositions 2 and 3. We proceed instead by studying the stochastic process driving a dynasty's (that is, a succession of generations') wealth accumulation.

To this end let $\lambda(t)$ denote a Poisson process with intensity p. Let then the stochastic process governing a dynasty's (that is, a series of successive generations') discounted

wealth, denoted $z(t)$ (with some abuse of notation), be defined by the following stochastic differential equation:

$$dz(t) = (g - g')dt - \xi(z(t^-))d\lambda(t)$$

where $\xi(z(t^-)) = \begin{cases} (b + \mu(1-b))z(t^-) & \text{with probability } \frac{p-q}{p} \\ \left(\frac{z(t^-)-w}{z(t^-)}\right)z(t^-) & \text{with probability } \frac{q}{p} \end{cases}$ and $z(t^-) = \lim_{s \to t^-} z(s)$.

The process $z(t)$ is a Markov jump process. Ergodicity of $z(t)$ implies the existence of an ergodic distribution of wealth in the population, $f(z)$. We shall study $z(t)$ by the Embedded Markov Chain method; see Borovkov (1998), ch. 4, and Karlin-Taylor (1981b), ch. 18.4. Let an embedded chain to $z(t)$ be constructed as follows. Let $z(t_n)$ denote the dynasty's discounted wealth at the end of the n'th generation. The sequence $\{z(t\_1), z(t\_2), \ldots\}$ defines a discrete time process, in fact a Markov chain, recursively by:

$$X_0 = z(0), \quad X_n = z(t_n)$$

where t_n is the stochastic process induced by the Poisson process $\lambda(t)$:[44]

$$\text{prob}(t_n = t) = \text{prob}(\lambda(t) = n \mid \lambda(\tau) < n, \ \forall \tau < t)$$

It is now straightforward to construct the transition kernel of the Markov chain X_n, denoted $P(X_n = x \mid X_{n-1} = z)$. Specifically,

$$P(X_n = x \mid X_{n-1} = z) = \begin{cases} q & \text{for } x = \underline{w} \\ 0 & \text{for } x \in (\underline{w}, (1-\mu)(1-b)z) \\ (p-q)\text{prob}(\lambda(\tau(x)) = n) & \text{for } x \in [(1-\mu)(1-b)z, \infty) \end{cases}$$

where $\tau(x)$ solves $x = ze^{(g-g')(\tau - t_{n-1})}(1-\mu)(1-b)$

---

[44]A related construction of an embedded Markov chain by means of a stochastic process t_n, in the context of Queue theory, is introduced in Karlin-Taylor (1981b), p. 497.

Ergodicity now follows by applying Theorem 13.3.3 of Meyn-Tweedie (1993), p. 323 if we show that the Markov chain X_n is positive Harris recurrent and aperiodic (we refer to Meyn-Tweedie (1993), ch. 5 and 10 for the precise definitions of these terms). In fact, these desired properties of X_n are straightforward to verify.

Proof of Proposition 5. Since

$$P = \frac{p - ((1-\mu)(1-b))^P q}{p - q((1-\mu)) - \tau} \tag{46}$$

The right hand side of (46) is increasing in P, for $P = 1$ it is larger than 1, and for $P \to \infty$ it is finite. Therefore we focus on the unique solution of $P \geq 1$. Since $p - q((1-\mu)) - \tau \geq 0$, and $0 \leq ((1-\mu)(1-b))^P \leq 1$ it follows that $lim_{\tau \to (p-q((1-\mu))} P = \infty$. Computing the derivatives of P, and substituting for $p - q((1-\mu)$ from 46 we get

$$\frac{dP}{d\tau} = \frac{P^2}{p - q\left\{((1-\mu)(1-b))^P (1 + P \ln((1-\mu)(1-b)))\right\}} > 0 \tag{47}$$

$$\frac{dP}{db} = \frac{P^2 ((1-\mu)(1-b))^{P-1} q (1-\mu)}{p - q\left\{((1-\mu)(1-b))^P (1 + P \ln((1-\mu)(1-b)))\right\}} \geq 0 \tag{48}$$

Note that $\frac{dP}{db} = 0$ would obtain only if $P \to \infty$ and $(1-\mu)(1-b) < 1$. As shown, this is indeed the case only if $\tau \to (p - q((1-\mu))$ and may also be ascertained directly by applying L'Hopital's rule to (48). To do this first apply L'Hopital's rule twice to $\frac{P^2}{((1-\mu)(1-b))^{-P+1}}$ and once to $\frac{P}{((1-\mu)(1-b))^P}$ by differentiating with respect to τ, and show that both expressions converge to zero as $\tau \to (p - q((1-\mu))$ because $lim_{\tau \to (p-q((1-\mu))} P = \infty$. Then substitute into the expression $\frac{dP}{db}$ to see that $lim_{\tau \to p-q((1-\mu)} \frac{dP}{db} = \frac{0}{p}$.

Derivation of the government budget constraint, (27). Consider the government expenditures (7), written in terms of discounted wealth and evaluated at the stationary distribution $f(z)$:

$$(p-q)\underline{w} + q \int_{\underline{w}}^{((1-b)(1-\mu))^{-1}\underline{w}} (\underline{w} - (1-b)(1-\mu)z)\, f(z) dz \tag{49}$$

Furthermore, the stationary distribution is

$$
\begin{aligned}
f(z) &= f\left(\underline{w}\right) \frac{z}{\underline{w}}^{-\left(\frac{p-a^*q((1-\mu)(1-b))}{(p-q((1-\mu))-\tau)}+1\right)} = \\
&= \frac{p-aq\left((1-\mu)(1-b)\right)}{(p-q((1-\mu))-\tau)} \underline{w}^{\frac{p-a^*q((1-\mu)(1-b))}{p-q((1-\mu))-\tau}} z^{-\left(\frac{p-a^*q((1-\mu)(1-b))}{p-q((1-\mu))-\tau}+1\right)}
\end{aligned}
$$

We proceed first by computing $\int_{\underline{w}}^{((1-b)(1-\mu))^{-1}\underline{w}} (1-b)(1-\mu)zf(z)dw$, obtaining

$$
(1-\mu)(1-b) \frac{p-aq\left((1-\mu)(1-b)\right)}{q(1-\mu)(1-a(1-b))+\tau}
$$
$$
\left(1-((1-\mu)(1-b))^{\frac{p-aq((1-\mu)(1-b))}{p-q((1-\mu))-\tau}-1}\right) \underline{w}
$$

Furthermore we compute $\int_{\underline{w}}^{((1-b)(1-\mu))^{-1}\underline{w}} f(z)dw$, obtaining

$$
\left(1-((1-\mu)(1-b))^{\frac{p-aq((1-\mu)(1-b))}{p-q((1-\mu))-\tau}}\right)
$$

Substituting the computations in (49), we conclude that government expenditures are:

$$
(p-q)\underline{w}+q\underline{w}\left(
\begin{array}{c}
\left(1-((1-\mu)(1-b))^{\frac{p-aq((1-\mu)(1-b))}{p-q((1-\mu))-\tau}-1}\right) \\
-(1-\mu)(1-b)\frac{p-aq(1-\mu)(1-b)}{q(1-\mu)(1-a(1-b))+\tau} \\
\left(1-((1-\mu)(1-b))^{\frac{p-aq(1-\mu)(1-b)}{p-q(1-\mu))-\tau}-1}\right)
\end{array}
\right)
$$

and therefore that the government budget constraint can be written as:

$$
\underline{w}\left(
\begin{array}{c}
(p-q)+q\left(1-((1-\mu)(1-b))^{\frac{p-aq((1-\mu)(1-b))}{p-q((1-\mu))-\tau}-1}\right) \\
\left(1-(1-\mu)(1-b)\frac{p-aq(1-\mu)(1-b)}{q(1-\mu)(1-a(1-b))+\tau}\right)
\end{array}
\right)
$$
$$
= (\tau+bq(1-\mu))M
$$

where M is average wealth.

Derivation of the discounted utility along the optimal path, (29). In our economy, the optimal consumption-savings path of an arbitrary agent is characterized by (3). Along this path, it is straightforward to compute

$$
U(z) = \int_t^\infty e^{(\theta+p)(t-\nu)}\left(\ln \eta w(t,\nu)+p\chi \ln(1-b)\chi\eta w(t,\nu)\right)d\nu
$$

where $w(t, \nu) = ze^{g(\nu - t)}$; or,

$$U(z) = \int_t^\infty e^{(\theta + p)(t - \nu)} \left(\ln \eta + \ln z + g(\nu - t) + p\chi \ln(1 - b)\chi\eta + p\chi \ln z + p\chi g(\nu - t) \right) d\nu$$

We proceed to analyze separately three components of $U(z)$:

i) $\int_t^\infty e^{(\theta + p)(t - \nu)} \left(\ln \eta + p\chi \ln(1 - b)\chi\eta \right) d\nu$;

ii) $\int_t^\infty e^{(\theta + p)(t - \nu)} \left(1 + p\chi \right) g(\nu - t) d\nu$; and finally,

iii) $\int_t^\infty e^{(\theta + p)(t - \nu)} \left(1 + p\chi \right) \ln z \, d\nu$.

i) Integrating,

$$\int_t^\infty e^{(\theta + p)(t - \nu)} \left(\ln \eta + p\chi \ln(1 - b)\chi\eta \right) d\nu = \frac{1}{\theta + p} \left(\ln \eta + p\chi \ln(1 - b)\chi\eta \right)$$

ii) Integrating by parts,

$$\int_t^\infty e^{(\theta + p)(t - \nu)} \left(1 + p\chi \right) g(\nu - t) d\nu =$$

$$= g\left(1 + p\chi \right) \left(-\frac{1}{\theta + p} \left[e^{(\theta + p)(t - \nu)}(\nu - t) \right]_t^\infty + \frac{1}{\theta + p} \int_t^\infty e^{(\theta + p)(t - \nu)} d\nu \right) =$$

$$= g\left(1 + p\chi \right) \left(-\frac{1}{\theta + p} \left[e^{(\theta + p)(t - \nu)(\nu - t)} \right]_t^\infty - \frac{1}{(\theta + p)^2} \left[e^{(\theta + p)(t - \nu)} \right]_t^\infty \right) =$$

$$= \frac{g\left(1 + p\chi \right)}{(\theta + p)^2}$$

iii) Integrating,

$$\int_t^\infty e^{(\theta + p)(t - \nu)} \left(1 + p\chi \right) \ln z \, d\nu = \frac{1 + p\chi}{\theta + p} \ln z$$

Adding up,

$$U(z) = \frac{1}{\theta + p} \left(\frac{g\left(1 + p\chi \right)}{\theta + p} + \ln \eta + p\chi \ln \left(1 - b \right)\chi\eta \right) + \frac{1 + p\chi}{\theta + p} \ln z$$

Derivatives of utilitarian social welfare with respect to b and τ.

$$\frac{\partial \Omega \left(M, b, \tau \right)}{\partial \tau} = \left(p + \theta \right)^{-2} \left(\left(1 + p\chi \right) \frac{P^{-2}}{P - 1} \frac{\partial P}{\partial \tau} - 1 \right) \tag{50}$$

$$= \left(p + \theta \right)^{-2} \left(\frac{\left(P - 1 \right)^{-1} \left(1 + p\chi \right)}{p - q \left\{ \left(\left(1 - \mu \right) \left(1 - b \right) \right)^{P} \left(1 + P \ln \left(\left(1 - \mu \right) \left(1 - b \right) \right) \right) \right\}} - 1 \right)$$

$$\frac{\partial \Omega \left(M, b, \tau \right)}{\partial b} \tag{51}$$

$$= \left(p + \theta \right)^{-1} \left(\frac{\left(1 + p\chi \right)}{p + \theta} \frac{P^{-2}}{P - 1} \frac{\partial P}{\partial b} - p\chi \left(1 - b \right)^{-1} \right)$$

$$= \left(p + \theta \right)^{-1} \left[\frac{\frac{\left(1 + p\chi \right)}{\left(P - 1 \right) \left(p + \theta \right)} \frac{\chi^{P-1} \left(p + \theta \right)^{P} \left(1 + p\chi \right)^{-1} \left(1 - b \right)^{P-1} q}{p - q \left\{ \left(\left(1 - \mu \right) \left(1 - b \right) \right)^{P} \left(1 + P \ln \left(\left(1 - \mu \right) \left(1 - b \right) \right) \right) \right\}}}{-p \left(1 - b \right)^{-1}} \right] \chi$$

where $\frac{\partial P}{\partial \tau}$ and $\frac{\partial P}{\partial b}$ are defined by (47) and (48) in the proof of Proposition 5 in Appendix A.

Appendix B: On the mechanisms possibly underlying a Pareto distribution of wealth

Various stochastic processes for individual wealth are known to aggregate into a Pareto distribution of wealth in the population; see Sornette (2000) for a technical review and Chipman (1976) for a careful and outstanding account of the historical contributions of this subject; see also Levy (2003).

One such process is exemplified here; its mathematical formulation first appears in Cantelli (1921).[45] Suppose a variable determining wealth (e.g., talent, age), which we denote α, is exponentially distributed. That is the number of people with $\alpha = \alpha_0$ is

$$N\left(\alpha_0\right) = pe^{-p\alpha_0}$$

Suppose wealth increases exponentially with α:

$$w = ae^{g\alpha}, \quad a > 0, \quad g \geq 0 \tag{52}$$

Therefore, we can solve for $\alpha = g^{-1}\ln\frac{w}{a}$, operate a change of variables and express the distribution of wealth as

$$N(w) = N\left(g^{-1}\ln\frac{w}{a}\right)\frac{d\alpha}{dw}$$

that is,

$$N(w) = \frac{p}{g}a^{\frac{p}{g}}w^{-\left(\frac{p}{g}+1\right)}$$

This is a Pareto distribution with the exponent $\frac{p}{g}$.[46]

---

[45] See also Fermi (1949)'s study of cosmic rays.

[46] A notable literature appeared in Italian in the first decades of the twentieth century which studies the wealth distribution resulting from different assumptions regarding the distribution of the generating factor we called α and on the functional dependence of wealth on this factor; see Chipman (1976) for a detailed discussion of these contributions.

The underlying mechanism which makes wealth Pareto distributed in our basic model is a similar one in which the factor α is represented by age. This is clearly illustrated by considering the simple economy with no bequests. At any time t, in this economy, the distribution of the population by age $t - s$ implied by the demographic structure of the economy is in fact

$$N\left(t - s\right) = p e^{-p(t-s)}$$

Moreover, abstracting from the complications of inheritance, each optimal consumption-savings choices imply a wealth accumulation process results in wealth increasing exponentially with age.

Appendix C: On the basic PDE and its solution by the *method of characteristics*

We illustrate in this Appendix the "method of characteristics" for the solution of partial differential equation (PDE's) by applying to a linear PDE with variable coefficients, a simple form of the PDE we solve in the paper. Consider the following PDE:

$$\frac{\partial f}{\partial t} = -af - bz\frac{\partial f}{\partial z} \tag{53}$$

with initial condition

$$f(z,0) = h(z)$$

Suppose first of all that the PDE is to be solved for $z \in \Re$, that is, that there is no boundary condition. The Method of Characteristics (see e.g., Farlow (1982), Ch. 27) requires solving the PDE in the characteristic space, (τ, t), implicitly constructed as follows:

$$\frac{dz}{d\tau} = bz, \quad \frac{dt}{d\tau} = 1$$

that is,

$$z(\tau) = c_1 e^{-b\tau}, \quad t(\tau) = \tau + c_2 \tag{54}$$

Let $z(0) = m$ and $t(0) = 0$, so that $c_1 = m$ and $c_2 = 0$. This construction has the property that the chain rule

$$\frac{df}{d\tau} = \frac{\partial f}{\partial z}\frac{\partial z}{\partial \tau} + \frac{\partial f}{\partial t}\frac{\partial t}{\partial \tau}$$

and (53) imply

$$\frac{df}{d\tau} = -af \tag{55}$$

a simple ordinary differential equation. The initial condition in characteristic space is $f(m,0) = h(m)$. The differential equation, together with the initial condition has solution

$$f(z(\tau), \tau) = h(m)e^{-a\tau}$$

Substituting back into the original space (z, t), using (54):

$$f(z, t) = h\left(ze^{-bt}\right) e^{-at} \tag{56}$$

In words: the density on z at time t is the same density that at time 0 was on ze^{-bt} dampened at a rate a.

Suppose now that the PDE is to be solved for $z \geq \underline{z}$, and that there is a boundary condition

$$f(\underline{z}, t) = B,$$

The Method of Characteristics applies to this class of problems, boundary value problems, as follows (see e.g., Hood (2003) and Strikwerda (2004), Ch. 1.2). The characteristic space is split along the characteristic $z = \underline{z}e^{b\tau}$. In particular, for $z \geq \underline{z}e^{b\tau}$ the solution to the PDE is determined by the initial condition, and

$$f(z, t) = h\left(ze^{-bt}\right) e^{-at}$$

For $z < \underline{z}e^{b\tau}$ the solution is instead determined by the boundary condition through the inverse transformation $\tau(z, y) = \ln \frac{z}{y} \frac{1}{b}$ and

$$f(z, t) = B\frac{1}{b}\left(\frac{z}{\underline{z}}\right)^{\frac{a}{b}}$$

Summarizing, the solution to the boundary value problem is:

$$f(z, t) = \begin{cases} B\frac{1}{b}\left(\frac{z}{\underline{z}}\right)^{\frac{a}{b}} & \text{for } z < \underline{z}e^{bt} \\ h\left(z\,e^{-bt}\right) e^{-at} & \text{for } z \geq \underline{z}e^{bt} \end{cases} \tag{57}$$

Appendix D: On Volterra-Fredholm integral equations of the second type

In this Appendix we report some results for the class of integral equations that we study in the paper. We consider Volterra-Fredholm integral equations of the second type with separable kernel:

$$f(z) = h(z) + \lambda \int_a^{\sigma(z)} K(y)H(z)f(y)dy$$

where the real maps h, σ, K, and H are continuously differentiable. It is convenient to study the following equivalent equation:

$$f(z) = h(z) + \lambda \int_a^{\infty} \tilde{K}(z,y)H(z)f(y)dy, \quad \tilde{K}(z,y) = K(y)I_{[a,\sigma(z)]}(z,y) \quad (58)$$

where $I_{[a,\sigma(z)]}(z,y)$ is the indicator function of the interval $[a,\sigma(z)]$,
$I_{[a,\sigma(z)]}(y) = \begin{cases} 1 & \text{for } y \in [a,\sigma(z)] \\ 0 & \text{otherwise} \end{cases}$. Note that $\tilde{K}(z,y)$ is not continuous. The theory of Volterra-Fredholm integral equations is, however, developed for square integrable kernels (see Tricomi (1957)), a condition which is obviously satisfied by $\tilde{K}(z,y)H(z)$. For the uniqueness of such solutions (excluding solutions that are zero almost everywhere) see Tricomi (1957), p. 10 and Chapter II and also p.63.

A simple explicit solution is reported by Polyanin-Manzhirov (1998), Ch. 2.1-7 (equation 50), for the following integral equation:

$$f(z) = h(z) + \lambda \int_a^z y^{\alpha_1} z^{\alpha_2} f(y)dy, \quad \text{for } \alpha_1 + \alpha_2 = -1$$

It corresponds to a special case of (58) in which:

$$\sigma(z) = z, \ \tilde{K}(z,y) = K(y) = y^{\alpha_1}, \ H(z) = z^{-\alpha_1 - 1}$$

Its solution is:

$$f(z) = h(z) + \int_a^z R(z, y)h(y)dy, \quad \text{for } R(z, y) = az^{\alpha_2 - \lambda}y^{\alpha_1 + \lambda}$$

Appendix E: On Earnings and Lump-sum Subsidies

Consider the economy in which each agent born at time t receives an expected present discounted value of earnings equal to $x(t)$. We assume that the present discounted value of earnings distributed at any time t in the economy, $x(t)$, is a constant fraction λ_x of aggregate wealth; that is, $x(t) = \lambda_x W(t)$. As a consequence, $x(t) = e^{g't}x$ for some initial value x.

In this case, as noted in the text, the total subsidy in the economy is:

$$(p-q)\left(\underline{w}(t) - x(t)\right) + q \int_{(\underline{w}(t)-x(t))}^{((1-b)(1-\mu))^{-1}(\underline{w}(t)-x(t))} \left(\underline{w}(t) - x(t) - (1-b)(1-\mu)w\right) f(w,t)dw$$

Moreover, the boundary condition is:

$$f\left(\underline{w}(t), t\right) = \frac{p-q}{g}\frac{1}{\underline{w}(t)} + \frac{1}{g}\frac{1}{\underline{w}(t)}q \int_{(\underline{w}(t)-x(t))}^{\sigma(\underline{w}(t)-x(t))} f(w,t)dw$$

The stationary distribution satisfies the integral equation (16). For this economy we also have

$$\sigma(y) = \frac{y-x}{(1-\mu)(1-b)}, \quad \text{and hence } \sigma(\underline{w}) = \underline{w}, \quad \frac{\partial \sigma(z)}{\partial z} = \frac{1}{(1-b)(1-\mu)}$$

We operate the transformation $j = \sigma(y) = \frac{y-x}{(1-\mu)(1-b)}$ and obtain, from (16):

$$\begin{aligned}
f\left(z\right) &= \left(\tfrac{z}{\underline{w}}\right)^{-\left(\frac{p}{g-g'}+1\right)} f\left(\underline{w}\right) \\
&+ q\left(g-g'\right)^{-1} \int_{\underline{w}}^{\frac{z-x}{(1-\mu)(1-b)}} f\left(j\right)\left[((1-\mu)(1-b)j + x)^{\left(\frac{p}{g-g'}\right)} (z)^{-\left(\frac{p}{g-g'}+1\right)}\right] dj
\end{aligned} \tag{59}$$

While we do not have of a closed form solution to this integral equation, a unique solution exists (see Appendix D). Moreover, Proposition (6) in the text states that, for large z, the distribution of discounted wealth is approximately Pareto. We provide here the proof of Proposition (6).

Proof of Prop. 6 The integral equation in this case, after the transformation $j = \sigma(y) = \frac{y-x}{(1-\mu)(1-b)}$ is reduced to:

$$f(z) = \left(\frac{z}{w}\right)^{-\left(\frac{p}{g-g'}+1\right)} f(w)$$
$$+q(g-g')^{-1} \int_{\underline{w}}^{\frac{z-x}{(1-\mu)(1-b)}} f(j) \left[((1-\mu)(1-b)j + x)^{\left(\frac{p}{g-g'}\right)} (z)^{-\left(\frac{p}{g-g'}+1\right)}\right] dj$$

A lower bound on $f(z)$, $l(z)$ is obtained by the solution to

$$l(z) = \left(\frac{z}{w}\right)^{-\left(\frac{p}{g-g'}+1\right)} l(w)$$
$$+q(g-g')^{-1} \int_{\underline{w}}^{w} f(j) \left[((1-\mu)(1-b)j + x)^{\left(\frac{p}{g-g'}\right)} (z)^{-\left(\frac{p}{g-g'}+1\right)}\right] dj =$$
$$= \left(\frac{z}{w}\right)^{-\left(\frac{p}{g-g'}+1\right)} f(w)$$

$l(z)$ is a power function with exponent $\frac{p}{p-q(1-\mu)-\tau}$, which is a Pareto distribution integrating to unity if defined over $w \geq f(\underline{w}) \left(\frac{p}{p-q(1-\mu)-\tau}\right)^{-1}$.

An upper bound on $f(z)$, $u(z)$ is obtained by the solution to

$$u(z) = \left(\frac{z}{w}\right)^{-\left(\frac{p}{g-g'}+1\right)} u(w)$$
$$+q(g-g')^{-1} \int_{\underline{w}}^{\frac{z}{(1-\mu)(1-b)}} u(j) \, j^{\frac{p}{g-g'}} z^{-\left(\frac{p}{g-g'}+1\right)} dj$$

since $1-\mu)(1-b)j + x \leq j$ by construction. Adapting the proof of Prop. 5, we can show that $u(z)$ is a power function with exponent

$$\frac{p - a^*q(1-\mu)(1-b)}{p - q(1-\mu) - \tau} > 1, \quad \text{for } 0 < a^* < 1 \text{ satisfying (23)}$$

which is a Pareto distribution integrating to one if defined for $w \geq f(\underline{w}) \left(\frac{p-a^*q(1-\mu)(1-b)}{p-q(1-\mu)-\tau}\right)^{-1}$. The distribution $f(z)$ for $z \geq f(\underline{w}) \left(\frac{p-a^*q(1-\mu)(1-b)}{p-q(1-\mu)-\tau}\right)^{-1}$ lies in between $u(z)$ and $l(z)$ [47], both

---

[47] Alternatively the solution of $f(z)$ may be explicitly written as the limiting solution obtained by the successive approximation method (see Polyanin and Manhzirov, section 9.9). The it is possible to show that the iterated kernels of $f(z)$ lie below the iterated kernels of $u(z)$.

of which converge to zero for large z and therefore it is approximated by

$$f(z) = \left(\frac{z}{\underline{w}}\right)^{-\left(\frac{p}{g-g'}+1\right)} f(\underline{w})$$
$$+q(g-g')^{-1} \int_{\underline{w}}^{\frac{z}{(1-\mu)(1-b)}} f(j) \left[((1-\mu)(1-b)j)^{\left(\frac{p}{g-g'}\right)} (z)^{-\left(\frac{p}{g-g'}+1\right)}\right] dj$$

which has the solution derived in the proof of Prop. 5: a Pareto distribution with exponent

$$\frac{p - a^* q (1-\mu)(1-b)}{p - q(1-\mu) - \tau} > 1, \quad \text{for } 0 < a^* < 1 \text{ satisfying (23)}$$

The analysis of the economy with earnings we just developed can be also applied to an economy without earnings but with a lump-sum subsidy rather than a means tested welfare policy. Consider the following welfare policy: all agents born at any t with an inheritance receive a lump-sum subsidy equal to $x(t)$ which grows at the aggregate economy's rate g': $x(t) = xe^{g't}$.[48]

In the case of lump-sum subsidies, the total amount of subsidies paid by the government at any time t is independent of the distribution of wealth at t and is a constant fraction of wealth at each time t:

$$(p - q)\underline{w} + qx$$

A fiscal policy (τ, b) determines the set of feasible welfare policies $(\underline{w}, x)$, which satisfies

$$(p - q)\underline{w}(t) + qx(t) = \tau W(t) + qb(1 - \mu)W(t)$$

We can now study the economy for which $0 < \mu, b < 1$ where welfare policies support a minimal discounted wealth $\underline{w}$ and provide all agents with discounted wealth greater that or equal to $\underline{w}$ with discounted lump-sum subsidies x. Under our assumptions it

---

[48] We assume for simplicity that

$$(1 - \mu)(1 - b)\underline{w}(t) + x(t) \geq \underline{w} \tag{60}$$

so that no inheriting agent has initial wealth smaller that the minimal wealth.

follows immediately that the boundary condition (19) holds, that is $f\left(\underline{w},t\right)=\frac{p-q}{g-g'}\frac{1}{\underline{w}}$. It is straightforward to notice that Proposition 6 can be applied to this case.

Appendix F: The Calibration - Sensitivity

The following tables report the Gini coefficient, G, for various combinations of b and τ:

$b\backslash\tau$	0	0.0025	0.0050	0.0075	0.0094
0	1.000	0.4432	0.1913	0.0661	0.0021
0.2	0.6896	0.3506	0.1728	0.0656	0.0021
0.4	0.5405	0.3079	0.1656	0.0655	0.0021
0.6	0.4674	0.2883	0.1630	0.0655	0.0021
0.8	0.4324	0.2801	0.1623	0.0655	0.0021
1	0.4201	0.2783	0.1622	0.0655	0.0021

and the ratio between minimum and average wealth, $\frac{w}{M}$,

$b\backslash\tau$	0	0.0025	0.0050	0.0075	0.0094
0	$UNDEF$	0.3858	0.6789	0.8760	0.9958
0.2	0.1837	0.4809	0.7053	0.8769	0.9958
0.4	0.2983	0.5291	0.7159	0.8771	0.9958
0.6	0.3629	0.5525	0.7197	0.8771	0.9958
0.8	0.3963	0.5623	0.7207	0.8771	0.9958
1	0.4083	0.5646	0.7208	0.8771	0.9958

The last table reports optimal taxes for different combinations of parameters $\frac{q}{p}$ and χ (the numbers in parenthesis are the optimal taxes - first τ and then b - while $\bar{\tau}=$ denotes the maximum feasible capital tax τ in the economy)

$\frac{q}{p} \backslash \chi$	8	10	12	15
.5	$(0.0126, 0)$	$(0.0119, 0)$	$(0.0112, 0)$	$(0.0103, 0)$
	$\bar{\tau} = 0.0128$	$\bar{\tau} = 0.0121$	$\bar{\tau} = 0.0114$	$\bar{\tau} = 0.0105$
.6	$(0.0119, 0)$	$(0.0111, 0)$	$(0.0103, 0)$	$(0.0092, 0)$
	$\bar{\tau} = 0.0121$	$\bar{\tau} = 0.0113$	$\bar{\tau} = 0.0105$	$\bar{\tau} = 0.0094$
.8	$(0.0107, 0)$	$(0.0096, 0)$	$(0.0085, 0)$	$(0.0071, 0)$
	$\bar{\tau} = 0.0109$	$\bar{\tau} = 0.0098$	$\bar{\tau} = 0.0087$	$\bar{\tau} = 0.0073$