A Theory of Optimal Inheritance Taxation*

Thomas Piketty, Paris School of Economics

Emmanuel Saez, UC Berkeley and NBER

November 20, 2012

Abstract

This paper derives optimal inheritance tax formulas that (a) capture the key equity-efficiency trade-off, (b) are expressed in terms of estimable sufficient statistics, (c) are robust to the underlying structure of preferences. We consider dynamic stochastic models with general and heterogeneous bequest tastes and labor productivities. We limit ourselves to simple but realistic linear or two-bracket tax structures to obtain tractable formulas. We show that long-run optimal inheritance tax rates can always be expressed in terms of distributional parameters, aggregate behavioral elasticities and social preferences for redistribution. Importantly, those results carry over with tractable modifications to (a) the case with social discounting (instead of steady-state welfare maximization), (b) the case with partly accidental bequests, (c) the standard Barro-Becker dynastic model. In all cases, the optimal inheritance tax rate increases with the concentration of bequest received and decreases with the elasticity of aggregate bequests to the net-of-tax rate. The optimal tax rate is positive and quantitatively large if concentration is high, the elasticity is low and society cares mostly about those receiving little inheritance. In contrast, the optimal tax rate is negative when society cares mostly about inheritors. We propose a calibration using micro-data for France and the United States. We find that for realistic parameters the optimal inheritance tax rate might be as large as 50%-60% - or even higher for top bequests, in line with historical experience.

*Thomas Piketty, Paris School of Economics, piketty@pse.ens.fr; Emmanuel Saez, University of California at Berkeley, Department of Economics, saez@econ.berkeley.edu. We are grateful to seminar participants at the Paris School of Economics, the London School of Economics, the National Bureau of Economic Research (Public Economics Program), the University of California at Berkeley, the Massachusetts Institute of Technology, and Stanford University for their comments. We thank Daron Acemoglu, Tony Atkinson, Alan Auerbach, Peter Diamond, Emmanuel Farhi, Mikhail Golosov, Louis Kaplow, Wojciech Kopczuk, Matt Weinzierl, Ivan Werning, and four anonymous referees for very helpful comments and stimulating discussions. We owe special thanks to Bertrand Garbinti for his help with the numerical calibrations. We acknowledge financial support from the Center for Equitable Growth at UC Berkeley. An earlier and longer draft was circulated as "A Theory of Optimal Capital Taxation," NBER Working Paper No. 17989, April 2012.

1 Introduction

There is substantial controversy both in the public policy debate and among economists about the proper level of taxation of capital income and inheritances. The public debate centers around the equity vs. efficiency issues. On the one hand, inequality between individuals arises not only from inequality in labor income but also from inequality in inheritances received. Piketty (2011) shows that bequests can be quantitatively important and tend to grow very rapidly in low-growth mature economies such as France. Furthermore, in contrast to labor income, individuals are hardly responsible for the inheritances they receive. Inheritance taxation redistributes from those receiving large inheritances toward those who do not. On the other hand, inheritance taxation also hurts those who accumulate wealth to leave inheritances for the welfare of their heirs, and hence might discourage wealth accumulation in the first place.

In the economic debate, there is a wide array of models and results capturing the issue of optimal capital/inheritance taxation. Those models differ primarily in terms of preferences for savings/bequests and the structure of economic shocks. In the infinite life model of Chamley (1986) and Judd (1985) with no stochastic shocks, the optimal capital income tax is zero in the long-run because a constant capital income tax creates a growing distortion on inter-temporal choices. In a life-cycle model with no bequests and homogeneous preferences for savings, labor income is the sole source of inequality and optimal capital income taxation is zero because a nonlinear earnings income tax is a more efficient tool for redistribution (Atkinson and Stiglitz, 1976). However, many subsequent studies have shown that those famous zero capital tax results can be overturned by relaxing each of the key hypotheses.[1] While those extensions provide important insights, it is hopeless to be able to ever measure directly the exact individual preferences' distributions to tell apart the different models. This leaves the current field of optimal capital income taxation fairly scattered with no clear policy implications as different–yet difficult to test–assumptions for bequest behavior lead to different formulas and magnitudes.

In this paper, we make progress on this issue by showing that optimal inheritance tax formulas can be expressed in terms of estimable "sufficient statistics" including distributional parameters, behavioral elasticities and social preferences for redistribution. Those formulas are robust to the underlying primitives of the model and capture the key equity-efficiency trade-off

[1] The most studied extensions leading to non-zero capital income taxes are: (a) presence of idiosyncratic labor income shocks, (b) accidental bequests, (c) bequests givers caring about pre-tax or post-tax bequests rather than the utility of heirs, (d) long-run steady-state welfare maximization, (e) time-invariant taxes, (f) lack of government commitment (Kopczuk 2013 and Diamond and Werning 2013 provide recent surveys).

in a transparent way. This approach has been fruitfully used in the analysis of optimal labor income taxation (Piketty and Saez, 2013 provide a recent survey). We follow a similar route and show that the same equity-efficiency trade-off logic also applies for inheritance taxation. This approach successfully brings together many of the main disparate results obtained in the literature on optimal capital tax theory.

We first consider in Section 2 dynamic stochastic models with general and heterogeneous preferences for bequests and ability for work where donors care solely about the net-of-tax bequest they leave to their heirs, and where the planner maximizes long-run steady-state welfare. This is the simplest case to illustrate the key trade-off transparently. Importantly however, we show in Section 3 that our results carry over with tractable modifications to (a) the case with social discounting (instead of steady-state welfare maximization), (b) the case with partly accidental bequests, (c) the standard Barro-Becker dynastic model with homogeneous discounting.

In all cases, the problem can be seen as an equity-efficiency trade-off where the optimal inheritance tax rate increases with the concentration of bequests received, decreases with the elasticity of aggregate bequests to the net-of-tax rate, and decreases with the value that society puts on the marginal consumption of bequest receivers and bequest leavers. The optimal tax rate is positive and quantitatively large if the elasticity is low, bequests are quantitatively large and highly concentrated, and society cares mostly about those receiving little inheritance. In contrast, the optimal tax rate is negative when society cares mostly about inheritors.

As in the public debate, the desirability of taxing bequests hinges primarily on wealth inequality and mobility and how social marginal welfare weights are distributed across groups. The optimal tax rate is zero when the elasticity of bequests is infinite nesting the zero capital tax Chamley-Judd result.[2] In contrast to the Atkinson-Stiglitz zero capital tax result in the life-cycle zero-bequest model, inheritance taxation is non-zero even with optimal labor taxation because with bequests, inequality is bi-dimensional ands labor income is no longer the unique determinant of lifetime resources. Indeed, in the OLG zero-bequest version of our dynamic model, labor income and capital income are perfectly correlated, inequality is uni-dimensional and the natural optimum capital income tax rate is zero under standard assumptions for preferences. The case with social discounting and government debt shows that cross-sectional redistribution

[2] In the altruistic dynasty case, the elasticity includes an anticipatory component as wealth accumulation decisions are affected before a tax change takes place. In the Chamley-Judd case with no stochastic shocks, this anticipatory elasticity is always infinite, even when preferences are tweaked (with endogenous discount rates) to make the long-run steady-state elasticity of bequests with respect to the long-run net-of-tax rate finite.

issues are orthogonal to optimal capital accumulation issues.

Importantly, we limit ourselves to extremely simple linear (or two-bracket) tax structures on inheritances and labor income to be able to obtain tractable formulas in models with very heterogeneous preferences. The advantages are that, by necessity, our tax system is well within the realm of current practice and the economic trade-offs appear transparently. This "simple tax structure" approach is in contrast to the recent new dynamic public finance (NDPF) literature (Kocherlakota, 2010 provides a recent survey) which considers the fully optimal mechanism given the informational structure. The resulting tax systems are complex–even with strong homogeneity assumptions–but potentially more powerful to increase welfare. Therefore, we view our approach as complementary to the NDPF approach.[3]

As an illustration of their use for policy recommendations, we propose in Section 4 a numerical simulation calibrated using micro-data for the case of France and the United States. We find that for realistic parameters the optimal inheritance tax rate might be as large as 50%-60% - or even higher for top bequests, in line with historical experience. To our knowledge, this is the first time that a model of optimal taxation delivers tractable and estimable formulas that can be used to analyze real world inheritance tax rates.

2 Optimal Steady-State Inheritance Tax Rate Formula

We consider a dynamic economy with a discrete set of generations $0, 1, .., t, ..$ and no exogenous growth. Each generation has measure one, lives one period, and is replaced by the next generation. Individual ti (from dynasty i living in generation t) receives pre-tax inheritance $b_{ti} \geq 0$ from generation $t-1$ at the beginning of period t. Inheritances earn an exogenous gross rate of return R per generation. Individual ti has exogenous pre-tax wage rate w_{ti}, works l_{ti}, and earns $y_{Lti} = w_{ti} l_{ti}$ at the end of period and then splits lifetime resources (the sum of net-of-tax labor income and capitalized bequests received) into consumption c_{ti} and bequests left $b_{t+1i} \geq 0$. We assume that there is a linear labor tax at rate τ_{Lt}, a linear tax on capitalized bequests at rate τ_{Bt}, and a lumpsum demogrant E_t.[4] Individual ti has utility function $V^{ti}(c, \underline{b}, l)$

[3] Farhi and Werning (2010) analyze optimal bequest taxation in a two-period Atkinson-Stiglitz model and in a NDPF model. In both cases the benchmark is zero bequest taxes and they show that caring about inheritors leads to negative but progressive bequest tax rates. In appendix A.4 we analyze the links between our analysis and theirs and show that their results have simple counter-parts within our linear tax structure analysis. We also discuss in which ways more complex tax mechanisms could in practice improve welfare.

[4] Note that the τ_{Bt} taxes both the raw bequest b_{t+1i} and the lifetime return to bequest $(R-1) \cdot b_{t+1i}$, so it should really be interpreted as a broad-based capital tax rather than as a narrow inheritance tax.

increasing in consumption $c = c_{ti}$ and net-of-tax capitalized bequests left $\underline{b} = Rb_{t+1i}(1-\tau_{Bt+1})$, and decreasing in labor supply $l = l_{ti}$. Hence, the individual maximization problem is:

$$\max_{l_{ti},c_{ti},b_{t+1i}\geq 0} V^{ti}(c_{ti}, Rb_{t+1i}(1-\tau_{Bt+1}), l_{ti}) \text{ s.t. } c_{ti} + b_{t+1i} = Rb_{ti}(1-\tau_{Bt}) + w_{ti}l_{ti}(1-\tau_{Lt}) + E_t. \quad (1)$$

The individual first order condition for bequests left is $V_c^{ti} = R(1-\tau_{Bt+1})V_{\underline{b}}^{ti}$ if $b_{t+1i} > 0$.

We denote by b_t, c_t, y_{Lt} aggregate bequests received, consumption, and labor income in generation t. We assume a stochastic process for utility functions V^{ti} and for wage rates w_{ti} such that, with constant tax rates and demogrant, the economy converges to a unique ergodic steady-state.[5] In the long-run, the position of each dynasty is independent of the initial position. Our model allows for heterogeneity in preferences for work and bequests and work ability.

We assume that the government chooses steady-state long-run policy E, τ_L, τ_B to maximize steady-state social welfare, a weighted sum of individual utilities with Pareto weights ω_{ti}, subject to a period-by-period budget balance $E_t = \tau_{Bt}Rb_t + \tau_{Lt}y_{Lt}$.[6]

$$SWF = \max_{\tau_L,\tau_B} \int_i \omega_{ti}V^{ti}(Rb_{ti}(1-\tau_B) + w_{ti}l_{ti}(1-\tau_L) + E - b_{t+1i}, Rb_{t+1i}(1-\tau_B), l_{ti}). \quad (2)$$

Note that, in the ergodic equilibrium, social welfare is constant over time. Taking the demogrant E as fixed, τ_L and τ_B are linked to meet the budget constraint, $E = \tau_B Rb_t + \tau_L y_{Lt}$. The aggregate variable b_t is a function of $1-\tau_B$ (assuming that τ_L adjusts), and y_{Lt} is a function of $1-\tau_L$ (assuming that τ_B adjusts). Formally, we can define the corresponding long-run elasticities as:

$$e_B = \frac{1-\tau_B}{b_t}\frac{db_t}{d(1-\tau_B)}\bigg|_E \quad \text{and} \quad e_L = \frac{1-\tau_L}{y_{Lt}}\frac{dy_{Lt}}{d(1-\tau_L)}\bigg|_E. \quad (3)$$

That is, e_B is the long run elasticity of aggregate bequest flow (i.e. aggregate capital accumulation) with repsect to the net-of-bequest-tax rate, while e_L is the long run elasticity of aggregate labor supply with repsect to the net-of-labor-tax rate.

We denote by $g_{ti} = \omega_{ti}V_c^{ti} / \int_j \omega_{tj}V_c^{tj}$ the normalized social marginal welfare weight on individual ti. g_{ti} measures the social value of increasing consumption of individual ti by \$1 (relative to distributing the \$1 equally across all individuals).

Consider a small reform $d\tau_B > 0$, budget balance with $dE = 0$ requires that $d\tau_L$ is such that:

$$Rb_t d\tau_B \left(1 - e_B\frac{\tau_B}{1-\tau_B}\right) = -d\tau_L y_{Lt}\left(1 - e_L\frac{\tau_L}{1-\tau_L}\right). \quad (4)$$

---

[5] All we need to assume is an ergodicity condition for the stochastic process for V^{ti} and w_{ti}: Whatever one's parental taste or productivity, there is always a positive probability of having any other taste or productivity. See Piketty and Saez (2012) for a precise mathematical statement and concrete examples. Random taste shocks typically generate Pareto distributions with realistic levels of wealth concentration–which are difficult to generate with labor productivity shocks alone. Random shocks to rates of return would work as well.

[6] Exogenous (non transfer) government spending H_t can be added without affecting the analysis.

Using the fact that b_{t+1i} and l_{ti} are chosen to maximize individual utility and applying the envelope theorem, the effect of the reform $d\tau_B, d\tau_L$ on steady-state social welfare is:

$$dSWF = \int_i \omega_{ti} V_c^{ti} \cdot (Rdb_{ti}(1-\tau_B) - Rb_{ti}d\tau_B - d\tau_L y_{Lti}) + \omega_{ti} V_{\underline{b}}^{ti} \cdot (-d\tau_B R b_{t+1i}).$$

At the optimum, $dSWF = 0$. Using the individual first order condition $V_c^{ti} = R(1-\tau_B)V_{\underline{b}}^{ti}$ when $b_{t+1i} > 0$, expression (4) for $d\tau_L$, and the definition of g_{ti}, we have:

$$0 = \int_i g_{ti} \cdot \left(-d\tau_B R b_{ti}(1+e_{Bti}) + \frac{1 - e_B\tau_B/(1-\tau_B)}{1 - e_L\tau_L/(1-\tau_L)} \frac{y_{Lti}}{y_{Lt}} R b_t d\tau_B - d\tau_B \frac{b_{t+1i}}{1-\tau_B} \right), \tag{5}$$

with $e_{Bti} = \frac{1-\tau_B}{b_{ti}} \frac{db_{ti}}{d(1-\tau_B)}\Big|_E$ the individual elasticity of bequest received.[7]

The first term captures the negative effect of $d\tau_B$ on bequest received (the direct effect and the dynamic effect via reduced pre-tax bequests), the second term captures the positive effect of reduced labor income tax, the third term captures the negative effect on bequest leavers.

To capture *distributional parameters* of earnings, bequests received, and bequests left, we use the ratios–denoted with an upper bar–of the average weighted by social marginal welfare weights g_{ti} to the unweighted average. Hence, $\bar{y}_L$, $\bar{b}^{\text{received}}$, $\bar{b}^{\text{left}}$, are the population averages of $g_{ti} \cdot y_{Lti}/y_{Lt}$, $g_{ti} \cdot b_{ti}/b_t$, $g_{ti} \cdot b_{t+1i}/b_{t+1}$. Such ratios are below one if the variable is lower for those with high social marginal welfare weights. Finally, let $\hat{e}_B$ be the average of e_{Bti} weighted by $g_{ti}b_{ti}$.[8] Dividing (5) by $Rb_t d\tau_B$, the first order condition is rewritten as:

$$0 = -\bar{b}^{\text{received}}(1+\hat{e}_B) + \frac{1 - e_B\tau_B/(1-\tau_B)}{1 - e_L\tau_L/(1-\tau_L)} \bar{y}_L - \frac{\bar{b}^{\text{left}}}{R(1-\tau_B)}, \quad \text{hence}$$

Steady-State Optimum. *The optimal tax rate τ_B that maximizes long-run steady state social welfare with period-by-period budget balance is given by:*

$$\tau_B = \frac{1 - \left[1 - \frac{e_L\tau_L}{1-\tau_L}\right] \cdot \left[\frac{\bar{b}^{\text{received}}}{\bar{y}_L}(1+\hat{e}_B) + \frac{1}{R}\frac{\bar{b}^{\text{left}}}{\bar{y}_L}\right]}{1 + e_B - \left[1 - \frac{e_L\tau_L}{1-\tau_L}\right] \frac{\bar{b}^{\text{received}}}{\bar{y}_L}(1+\hat{e}_B)}. \tag{6}$$

Five important points are worth noting about this formula.

First, the presence of R in formula (6) is a consequence of considering steady-state maximization, i.e., no social discounting. As shown in Section 3.1, with social discounting at rate $\Delta < 1$, R should be replaced by $R\Delta$. Furthermore, in a closed economy with government debt, dynamic efficiency implies that the modified Golden, $R\Delta = 1$, holds. Hence, formula (6) continues to apply in the canonical case with discounting and dynamic efficiency by replacing R by

[7]Note that e_B is the bequest-weighted population average of e_{Bti}.

[8]$\hat{e}_B$ is equal to e_B (b_{ti}-weighted average of e_{Bti}) if individual bequest elasticities are uncorrelated with g_{ti}.

one in equation (6). This also remains true with exogenous economic growth.[9] Therefore, if one believes that the natural benchmark case involves dynamic efficiency and no social discounting ($\Delta = 1$), then it is natural to interpret formula (6) assuming $R = 1$. It is unclear however whether this is the most relevant case for numerical calibrations, as we shall see below.

Second, τ_B decreases with the elasticity e_B and with $\bar{b}^{\text{received}}$ and $\bar{b}^{\text{left}}$, i.e., the social weight put on bequests receivers and leavers. Under a standard utilitarian case with decreasing marginal utility of income, welfare weights g_{ti} are low when when bequests and/or labor income are high. As bequests are more concentrated than labor income (Piketty, 2011), we expect $\bar{b}^{\text{received}} < \bar{y}_L$ and $\bar{b}^{\text{left}} < \bar{y}_L$. In the limit where bequests are infinitely concentrated, $\bar{b}^{\text{received}}, \bar{b}^{\text{left}} \ll \bar{y}_L$ and (6) boils down to $\tau_B = 1/(1 + e_B)$, the standard revenue maximizing rate. Conversely, when social welfare weights g_{ti} put sufficient importance on large inheritors, then $\bar{b}^{\text{received}}$ will be larger than one and τ_B will eventually become negative. If society cares mostly about very large inheritors, then the optimal τ_B will be infinitely negative.

Third and related, bequest taxation differs from capital taxation in a standard OLG model with no bequests in two ways. Firstly, τ_B hurts both donors and donees making bequests taxation relatively less desirable, as τ_B decreases with both $\bar{b}^{\text{left}}$ and $\bar{b}^{\text{received}}$ (Kaplow 2001). Indeed, a negative τ_B is possible when redistributive motives are moderate so that $\bar{b}^{\text{left}}, \bar{b}^{\text{received}}$, and $\bar{y}_L$ are all close to one and the labor supply elasticity e_L is modest. Secondly, bequests introduce a new dimension of life-time resources inequality lowering $\bar{b}^{\text{received}}/\bar{y}_L$, $\bar{b}^{\text{left}}/\bar{y}_L$ and making bequests taxation more desirable. This intuition is made precise in Section 3.2 where we consider the OLG model where zero capital taxation is the natural benchmark.

Fourth, general social marginal welfare weights allow great flexibility in the social welfare criterion choice. One normatively appealing concept is that individuals should be compensated for inequality they are not responsible for–such as bequests received–but not for inequality they are responsible for–such as labor income (Fleurbaey, 2008). This amounts to setting social welfare weights to zero for all bequest receivers and setting them positive and uniform on zero-bequests receivers. About half the population receives negligible bequests so that this "Meritocratic Rawlsian" optimum has broader appeal than the standard narrow Rawlsian case.

Meritocratic Rawlsian Steady-State Optimum. *The optimal tax rate τ_B that maximizes*

[9] We started with the steady-state, budget balance, open economy, no growth case for pedagogical reasons. Introducing social discounting, government debt, closed economy, and economic growth requires more complex notations and derivations even though the core of the proof and the key economic mechanisms are the same.

long-run welfare of zero-bequest receivers with period-by-period budget balance is given by:

$$\tau_B = \frac{1 - \left[1 - \frac{e_L \tau_L}{1-\tau_L}\right] \frac{1}{R} \frac{\bar{b}^{\text{left}}}{\bar{y}_L}}{1 + e_B}, \quad \textit{with} \quad \bar{b}^{\text{left}}, \bar{y}_L \tag{7}$$

the ratios of average bequests left, and labor income of zero–receivers to population averages.

In that case, even if zero-receivers have population average labor earnings ($\bar{y}_L = 1$), if bequests are quantitatively important in life-time resources, zero-receivers will leave smaller bequests than average so that $\bar{b}^{\text{left}} < 1$. Formula (7) then implies $\tau_B > 0$ even with $R = 1$ and $e_L = 0$. The optimum rate (7) is below the revenue maximizing rate $1/(1+e_B)$ as zero-bequests receivers are hurt by the inheritance tax through the bequests they leave to their heirs.

Fifth, the optimal τ_B decreases with e_L (and τ_L) as a more elastic labor supply makes it less desirable to shift bequest taxation to labor taxation. In the inelastic labor case, formula (7) further simplifies to $\tau_B = \frac{1-\bar{b}^{\text{left}}/(R\bar{y}_L)}{1+e_B}$. If we further assume $e_B = 0$ and $R = 1$ (benchmark case with dynamic efficiency and $\Delta = 1$), then the optimal tax formula becomes $\tau_B = 1 - \bar{b}^{\text{left}}/\bar{y}_L$. The optimal tax rate solely depends on distributional parameters, namely the relative position of zero-bequest receivers in the distributions of bequests left and labor income. For instance, if $\bar{b}^{\text{left}}/\bar{y}_L = 10\%$, e.g. if zero-bequest receivers expect to leave bequests that are only one tenth of average bequests and to receive the same average labor income, then it is in their interest to tax bequests at rate $\tau_B = 90\%$. Intuitively, with a 90% bequest tax rate, the distortion on the "bequest left" margin is so large that the utility value of one additional dollar devoted to bequests is 10 times larger than one additional dollar devoted to consumption.[10] This intuition illustrates the critical importance of distributional parameters–and also of perceptions about these parameters. If everybody expects to leave large bequests, then subjectively optimal bequest tax rates will unsurprisingly be fairly small–or even negative.

Naturally, as for virtually all optimal tax formulas, $\bar{b}^{\text{left}}, \bar{b}^{\text{received}}$, and $\bar{y}_L$ depend on τ_B, τ_L and hence are endogenous. Assumptions need to be made on how those parameters vary with tax rates. To demonstrate the practical power of those sufficient statistics formulas, we propose such a calibration exercise in Section 4 using the actual joint distributions of $(b^{\text{received}}, b^{\text{left}}, y_L)$ for the United States and France. As an aside, it is also possible to obtain a formula trading-off τ_L and the optimal lumpsum demogrant E.[11] In all cases, formula (6) applies unchanged.

[10] For the same reasons, note that if $\bar{b}^{\text{left}}/\bar{y}_L = 100\%$, but $R = 2$, then $\tau_B = 50\%$. I.e. if the return to capital doubles the value of bequests left at each generation, then it is in the interest of zero receivers to tax capitalized bequest at a 50% rate, even if they plan to leave as much bequests as the average.

[11] The optimal τ_L (or E) increases with the social redistributive tastes and decreases with the overall responsiveness of labor and bequests to taxation.

3 Extending the Optimal Formula to other Contexts

In this section, we show that the formula we obtained in the basic model of Section 2 can be easily adapted to alternative settings. The basic equity-efficiency trade-off carries over in all cases showing that the sufficient statistics approach is extremely powerful to bring together various strands of the optimal capital tax literature that previously seemed irreconcilable.

3.1 Social Discounting, Government Debt, and Dynamic Efficiency

Let us assume that the government maximizes a discounted stream of social welfare across periods with generational discount rate $\Delta \leq 1$ (Section 2 considered the special case $\Delta = 1$):

$$SWF = \sum_{t \geq 0} \Delta^t \int_i \omega_{ti} V^{ti}(Rb_{ti}(1-\tau_{Bt}) + w_{ti}l_{ti}(1-\tau_{Lt}) + E_t - b_{t+1i}, Rb_{t+1i}(1-\tau_{Bt+1}), l_{ti}).$$

Budget balance and open economy. Let us first keep period-by-period budget balance so that $E_t = \tau_{Bt}Rb_t + \tau_{Lt}y_{Lt}$ along with the open economy R exogenous assumption. Consider again a reform $d\tau_B$ so that $d\tau_{Bt} = d\tau_B$ for all $t \geq T$ (and correspondingly $d\tau_{Lt}$ to maintain budget balance and keeping E_t constant) with T large (so that all variables have converged).

$$dSWF = \sum_{t \geq T} \Delta^t \int_i \omega_{ti} V_c^{ti} \cdot (Rdb_{ti}(1-\tau_{Bt}) - Rb_{ti}d\tau_B - d\tau_{Lt}y_{Lti}) + \sum_{t \geq T-1} \Delta^t \int_i \omega_{ti} V_{\underline{b}}^{ti} \cdot (-d\tau_B Rb_{t+1i}).$$

The key difference with steady-state maximization is that the reform also hurts generation $T-1$ bequest leavers. Hence, the negative term on bequest leavers carries extra weight $1/\Delta$ in the social welfare calculus. Defining e_B, $\hat{e}_B$, e_L as the average discounted elasticities (see appendix A.1.1 for exact and complete definitions), we obtain the following optimal formula:

Long-run optimum with social discounting. *The optimal long-run tax rate τ_B that maximizes discounted social welfare with period-by-period budget balance is given by:*

$$\tau_B = \frac{1 - \left[1 - \frac{e_L\tau_L}{1-\tau_L}\right] \cdot \left[\frac{\bar{b}^{\text{received}}}{\bar{y}_L}(1+\hat{e}_B) + \frac{1}{R\Delta}\frac{\bar{b}^{\text{left}}}{\bar{y}_L}\right]}{1 + e_B - \left[1 - \frac{e_L\tau_L}{1-\tau_L}\right] \frac{\bar{b}^{\text{received}}}{\bar{y}_L}(1+\hat{e}_B)}, \tag{8}$$

The only difference with (6) is that R is replaced by $R\Delta$ in the denominator of the term reflecting the utility loss of bequest leavers. The intuition is transparent: the utility loss of bequest leavers has a multiplicative factor $1/\Delta$ because bequests leavers are hurt one generation in advance of the tax reform.[12] Naturally, with no discounting $\Delta = 1$ and formulas (6) and (8) coincide.

[12] A future inheritance tax increase 30 years away, does not generate any revenue for 30 years and yet already hurts the current adult population contemplating leaving bequests to their heirs in 30 or more years.

Government debt in the closed economy. Suppose now that the government can use debt (paying the same rate of return R) and hence can transfer resources across generations. Let a_t be the net asset position of the government. If $R\Delta > 1$, reducing consumption of generation t to increase consumption of generation $t+1$ is desirable (and vice-versa). Hence, if $R\Delta > 1$, the government wants to accumulate infinite assets. If $R\Delta < 1$, the government wants to accumulate infinite debts. In both cases, the small open economy assumption would cease to hold. Hence, a steady-state equilibrium only exists if the *Modified Golden Rule* $R\Delta = 1$ holds.

Therefore, it is natural to consider the closed-economy case with endogenous capital stock $K_t = b_t + a_t$, CRS production function $F(K_t, L_t)$, where L_t is the total labor supply, and rates of returns on capital and labor are given by $R_t = 1 + F_K$ and $w_t = F_L$. Denoting by $\underline{R}_t = R_t(1-\tau_{Bt})$ and $\underline{w}_t = w_t(1-\tau_{Lt})$ the after-tax factor prices, the government budget dynamics is given by $a_{t+1} = R_t a_t + (R_t - \underline{R}_t) b_t + (w_t - \underline{w}_t) L_t - E_t$. Two results can be obtained in that context.

First, going back for an instant to the budget balance case, it is straightforward to show that formula (8) carries over unchanged in this case. This is a consequence of the standard optimal tax result of Diamond and Mirrlees (1971) that optimal tax formulas are the same with fixed prices and endogenous prices. The important point is that the elasticities e_B and e_L are pure supply elasticities (i.e., keeping factor prices constant). Intuitively, the government chooses the net-of-tax prices $\underline{R}_t$ and $\underline{w}_t$ and the resource constraint is $0 = b_t + F(b_t, L_t) - \underline{R}_t b_t - \underline{w}_t L_t - E_t$ so that the pre-tax factors effectively drop out of the maximization problem and the same proof goes through (see appendix A.1.2). Second, and most important, moving to the case with debt, we can show that the long-run optimum takes the following form.

Long-run optimum with social discounting, closed economy, and government debt. *In the long-run optimum, the* ***Modified Golden Rule*** *holds so that* $R\Delta = 1$. *The optimal long-run tax rate* τ_B *continues to be given by formula (8) with* $R\Delta = 1$,

$$\tau_B = \frac{1 - \left[1 - \frac{e_L \tau_L}{1-\tau_L}\right] \cdot \left[\frac{\bar{b}^{\text{received}}}{\bar{y}_L}(1+\hat{e}_B) + \frac{\bar{b}^{\text{left}}}{\bar{y}_L}\right]}{1 + e_B - \left[1 - \frac{e_L \tau_L}{1-\tau_L}\right] \frac{\bar{b}^{\text{received}}}{\bar{y}_L}(1+\hat{e}_B)}. \tag{9}$$

Proof: We first establish that the Modified Golden Rule holds in the long-run. Consider a small reform $d\underline{w}_T = d\underline{w} > 0$ for a single T large (so that all variables have converged). Such a reform has an effect $dSWF$ on discounted social welfare (measured as of period T) and da on long-term government debt (measured as of period T). Both $dSWF$ and da are proportional to dw.

Now consider a second reform $d\underline{w}_{T+1} = -Rd\underline{w} < 0$ at $T+1$ only. By linearity of small changes, this reform has welfare effect $dSWF' = -R\Delta dSWF$ as it is $-R$ times larger and happens one period after than the first reform. The effect on government debt is $da' = -Rda$ measured as of period $T+1$, and hence $-da$ measured as of period T (i.e., the same absolute effect as the initial reform). Hence, the sum of the two reforms would be neutral for government debt. Therefore, if social welfare is maximized, the sum has to be neutral from a social welfare perspective as well implying that $dSWF + dSWF' = 0$ so that $R\Delta = 1$.

Next, we can easily extend the result above that the optimal tax formula takes the same form with endogenous factor prices (appendix A.1.2). Hence, (8) applies with $R\Delta = 1$. *Q.E.D.*

This result shows that dynamic efficiency considerations (i.e., optimal capital accumulation) are conceptually orthogonal to cross-sectional redistribution considerations. That is, whether or not dynamic efficiency prevails, there are distributional reasons pushing for inheritance taxation, as well as distortionary effects pushing in the other direction, resulting into an equity-efficiency trade-off that is largely independent from aggregate capital accumulation issues.[13]

One natural benchmark would be to assume that we are at the Modified Golden Rule (though this is not necessarily realistic). In that case, the optimal tax formula is independent of R and Δ and depends solely on elasticities and the redistributive factors $\bar{b}^{\text{received}}, \bar{b}^{\text{left}}, \bar{y}_L$.

If the Modified Golden Rule does not hold (which is probably more plausible) and there is too little capital so that $R\Delta > 1$, then the welfare cost of taxing bequests left is smaller and the optimal tax rate on bequests should be higher (everything else being equal). The intuition for this result is simple: if $R\Delta > 1$, pushing resources toward the future is desirable. Taxing bequests more in period T hurts period $T-1$ bequest leavers and benefits period T labor earners, effectively creating a transfer from period $T-1$ toward period T. This result and intuition depend on our assumption that bequests left by generation $t-1$ are taxed in period t as part of generation t life-time resources. This fits with actual practice as bequests taxes are paid by definition at the end of the lives of bequests leavers and paid roughly in the middle of the adult life of bequests receivers.[14] If we assume instead that period t taxes are $\tau_{Bt} b_{t+1} + \tau_{Lt} y_{Lt}$, then formula (8) would have no $R\Delta$ term dividing $\bar{b}^{\text{left}}$ but all the terms in $\bar{b}^{\text{received}}$ would be multiplied by $R\Delta$. With government debt and dynamic efficiency ($R\Delta = 1$), formula (9) no

[13] The same decoupling results have been proved in the OLG model with only life-cycle savings with linear Ramsey taxation and a representative agent per generation (King, 1980 and Atkinson and Sandmo, 1980).

[14] Piketty and Saez (2012) make this point formally with a continuum of overlapping cohorts. With accounting budget balance, increasing bequests taxes today allows to reduce labor taxes today, hurting the old who are leaving bequests and benefiting current younger labor earners (it is too late to reduce the labor taxes of the old).

longer depends on the timing of tax payments.[15]

Economic growth. Normatively, there is no good justification for discounting the welfare of future generations, i.e. for assuming $\Delta < 1$. However, with $\Delta = 1$, the Modified Golden Rule implies that $R = 1$ so that the capital stock should be infinite. A standard way to eliminate this unappealing result as well as making the model more realistic is to consider standard labor augmenting economic growth at rate $G > 1$ per generation. Obtaining a steady-state where all variables grow at rate G per generation requires imposing standard homogeneity assumptions on individual utilities so that $V^{ti}(c, \underline{b}, l) = \frac{\left(U^{ti}(c,\underline{b})e^{-h_{ti}(l)}\right)^{1-\gamma}}{1-\gamma}$ with $U^{ti}(c, \underline{b})$ homogeneous of degree one. In that case, labor supply is unaffected by growth. The risk aversion parameter γ reflects social redistributive tastes both within and across generations.[16] We show in appendix A.1.3 the following results with economic growth.

First, the steady-state optimum formula (6) carries over by just replacing R by R/G. The intuition is simple. Leaving a relative bequest b_{t+1i}/b_{t+1} requires making a bequest G times larger than leaving the same relative bequest b_{t+1i}/b_t. Hence, the relative cost of taxation to bequest leavers is multiplied by a factor G.

Second, with social discounting at rate Δ, marginal utility of consumption grows at rate $G^{-\gamma} < 1$ as future generations are better off and all macro-economic variables grow at rate G. This amounts to replacing Δ by $\Delta G^{1-\gamma}$ in the $dSWF$ calculus. Hence, with those two new effects, formula (8) carries over simply replacing ΔR by $\Delta(R/G)G^{1-\gamma} = \Delta R G^{-\gamma}$.

Third, with government debt in a closed economy, the Modified Golden Rule becomes $\Delta R G^{-\gamma} = 1$ (equivalent to $r = \delta + \gamma g$ when expressed in conventional net instantaneous returns). The well-known intuition is the following. One dollar of consumption in generation $t+1$ is worth $\Delta G^{-\gamma}$ dollars of consumption in generation t because of social discounting Δ and because marginal utility in generation $t+1$ is only $G^{-\gamma}$ times the marginal utility of generation t. At the dynamic optimum, this must equal the rate of return R on government debt. Importantly, with the Modified Golden Rule, formula (9) carries over unchanged with growth.[17]

Role of R and G. Which formula should be used? From a purely theoretical viewpoint, it

[15]In the Meritocratic Rawlsian optimum, we can obtain (9) by considering steady state maximization subject to $\tau_{Bt} b_{t+1} + \tau_{Lt} y_{Lt} = E_t$ and without the need to consider dynamic efficiency issues (see appendix A.3).

[16]In general, the private risk aversion parameter γ might well vary across individuals, and differ from the social redistributive taste Γ. Here we ignore this possibility so as to simplify the notations. See Piketty and Saez, 2012, Appendix C, for a more complete treatment.

[17]Adding population growth at rate N per generation does not affect formula (9) either. The Modified Golden rule is unchanged if social welfare each period grows linearly in population size (Benthamite case). The Modified Golden rule becomes $\Delta R G^{-\gamma}/N = 1$ if period-by-period social welfare does not grow with population size.

is more natural to replace R by $\Delta R G^{-\gamma} = 1$ in formula (6), so as to entirely separate the issue of optimal capital accumulation from that of optimal redistribution. In effect, optimal capital accumulation is equivalent to removing all returns to capital in the no growth model ($R = 1$). However from a practical policy viewpoint, it is probably more justified to replace R by R/G in formula (6) and to use observed R and G in order to calibrate the formula. The issue of optimal capital accumulation is very complex, and there are many good reasons why the Modified Golden Rule $\Delta R G^{-\gamma} = 1$ does not seem to be followed in the real world.[18] In practice, it is very difficult to know what the optimal level of capital accumulation really is. Maybe partly as a consequence, governments tend not to interfere too massively with the aggregate capital accumulation process and usually choose to let private forces deal with this complex issue (net government assets - positive or negative - are typically much smaller than net private assets). One pragmatic viewpoint is to take these reasons as given and impose period-by-period budget constraint (so that in effect the government does not interfere at all with aggregate capital accumulation), and consider steady-state maximization, in which case we obtain formula (6) with R/G.

Importantly, the return rate R and the growth rate G matter for optimal inheritance rates even in the case with dynamic efficiency. Namely, a larger R/G implies a higher level of aggregate bequest flows (Piketty 2011), and also a higher concentration of inherited wealth. As a result, a larger R/G tends to imply both smaller $\bar{b}^{\text{received}}$ and smaller $\bar{b}^{\text{left}}$ among bottom receivers and hence a higher τ_B.

3.2 Role of Bi-Dimensional Inequality: Contrast with the OLG Model

Our results on positive inheritance taxation (under specific redistributive social criteria) hinge crucially on the fact that, with inheritances, labor income is no longer a complete measure of life-time resources, i.e., that our model has bi-dimensional (labor income, inheritance) inequality.

To see this, consider a simple OLG model with no bequest and where each person lives for two periods and works only in period 1 and funds consumption in period 2 with life-cycle savings. In this model, inequality is uni-dimensional (solely due to labor income). Appendix A.2 shows that the OLG case fits within the class of economies we have considered by simply relabeling

[18] E.g. with $\Delta = 1$ but high social curvature γ (intergenerational Rawlsian), MGR implies that we should leave vanishingly small capital to future generations - they will be so productive that they will not need it. One obvious problem with this reasoning is endogenous growth. If little capital is left, then future growth might also be lower. See Piketty and Saez 2012 Appendix B for a more detailed discussion of dynamic efficiency issues, in particular in relation to the debates on environmental capital and global warming.

bequests as life-cycle savings and assuming that each dynasty is a succession of two-period long lives. The young work and receive no bequests. The old do not work and receive "bequests" from their younger self. As welfare of the old is captured through the young, the zero-receivers social objective is the natural benchmark. In this model, our optimal formulas apply and the optimal tax rate on savings is zero under standard homogeneity assumptions. Intuitively, the linear labor tax is sufficient to redistribute across the uni-dimensional labor income inequality.

Optimal zero capital tax in the no-bequest OLG version of our model. *If individual utilities are such that* $V^{ti}(c, \underline{b}, l) = U^{ti}(u(c, \underline{b}), l)$ *with* $u(c, \underline{b})$ *homogeneous of degree 1 and homogeneous in the population, linear labor taxation is sufficient and zero savings tax is optimal.*

The formal proof is presented in appendix A.2 where we show that any tax system (τ_B, τ_L, E) can be replaced by a tax system $(\tau'_B = 0, \tau'_L, E')$ that leaves everybody as well off and raises more revenue. This result is the linear tax version of the Atkinson-Stiglitz theorem. In addition to standard weak separability, it also requires that $u(c, \underline{b})$ be homogeneous of degree one (Deaton, 1979).[19] The intuition can be understood using our optimal formula (9) for zero receivers. Suppose for simplicity here that there is no demogrant. With $u(c, \underline{b})$ homogeneous, bequest decisions are linear in life-time resources so that $b_{t+1i} = s \cdot y_{Lti}(1 - \tau_{Lt})$ where s is homogeneous in the population but may depend on the price $1/[R(1 - \tau_{Bt+1})]$. This immediately implies that $E[\omega_{ti} V_c^{ti} b_{t+1i}]/b_{t+1} = E[\omega_{ti} V_c^{ti} y_{Lti}]/y_{Lt}$ so that $\bar{b}^{\text{left}} = \bar{y}_L$. Absent any behavioral response, capital taxes are equivalent to labor taxes on distributional grounds in the OLG model because there is only one dimension of inequality left. Next, the capital tax τ_B also reduces labor supply (as it reduces the use of income) exactly in the same proportion as the labor tax. Hence, shifting from the labor tax to the capital tax has zero net effect on labor supply and $e_L = 0$. With $\bar{b}^{\text{left}} = \bar{y}_L$ and $e_L = 0$ (and $\bar{b}^{\text{received}} = 0$), (9) implies that $\tau_B = 0$. Critically, this zero-tax result and reasoning fails with bequests as they create a second dimension of inequality and $\bar{b}^{\text{left}} < \bar{y}_L$.

An extension would be to consider nonlinear (but static) earnings taxation. With inheritances, the Atkinson-Stiglitz theorem would no longer apply as, conditional on labor earnings, bequests left are a signal for bequests received, and hence correlated with social marginal welfare weights violating assumption 1 of Saez (2002) extension of Atkinson-Stiglitz to heterogeneous populations.[20] We leave the derivation of the optimal τ_B in that case to future work.[21]

[19]With economic growth, this homogeneity assumption is also needed to obtain balanced growth (Section 3.1).

[20]The simplest way to see this is to consider the case with uniform labor earnings. As inequality arises solely from bequests, labor taxation is useless for redistribution and bequest taxation is the only redistributive tool.

[21]As shown in appendix A.4, this OLG model is equivalent to the Farhi and Werning (2010) two-period model with working parents leaving bequests and non-working children receiving bequests. When children's welfare

3.3 Accidental Bequests or Wealth Lovers

Individuals also leave bequests for non-altruistic reasons. For example, some individuals may value wealth per se (e.g., it brings social prestige and power), or for precautionary motives and leave accidental bequests due to imperfect annuitization. Such non-altruistic reasons are quantitatively important (Kopczuk and Lupton, 2007). If individuals do not care about the after-tax bequests they leave, they are not hurt by bequests taxes on bequests they leave.[22] Bequest receivers continue to be hurt by bequests taxes. This implies that the last term $\bar{b}^{\text{left}}$ in the numerator of our formulas capturing the negative effect of τ_B on bequest leavers ought to be discounted. Formally, it is straightforward to generalize the model to utility functions $V^{ti}(c, b, \underline{b}, l)$ where b is pre-tax bequest left which captures wealth loving motives. The individual first order condition becomes $V_c^{ti} = R(1-\tau_{Bt+1})V_{\underline{b}}^{ti} + V_b^{ti}$ and $\nu_{ti} = R(1-\tau_{Bt+1})V_{\underline{b}}^{ti}/V_c^{ti}$ naturally captures the relative importance of altruism in bequests motives. All our formulas carry over by simply redefining $\bar{b}^{\text{left}}$ as the population average of $g_{ti}\nu_{ti}b_{ti}/b_t$ instead of $g_{ti}b_{ti}/b_t$. As we shall see in Section 4, existing surveys can be used to measure the relative importance of altruistic motives vs. other motives to calibrate the optimal τ_B. Hence, our approach is robust and flexible to accommodate the non altruistic motive extension that is empirically first order.

3.4 Standard Dynastic Model

The Barro-Becker dynastic model has been widely used in the analysis of optimal capital taxation. Our sufficient statistics formula approach can also fruitfully be used in that case with minor modifications. In the dynastic model, individuals care about the utility of their heirs V^{t+1i} instead of the after-tax capitalized bequests $R(1-\tau_{Bt+1})b_{t+1i}$ they leave. The standard assumption is the recursive additive form $V^{ti} = u^{ti}(c, l) + \delta V^{t+1i}$ where $\delta < 1$ is a uniform discount factor. With stochastic wages and standard assumptions, this model also generates an ergodic equilibrium where long-run individual outcomes are independent of initial position. The key difference with our initial model is that bequest behavior can change generations in advance of an anticipated tax change.[23]

To illustrate our approach in the dynastic model in the most pedagogical way, let us first

counts solely through parents' altruism for social welfare, the zero-receivers formula applies and the optimal τ_B is zero as just discussed. If additional weight is put on children's welfare, the formula with $\bar{b}^{\text{received}} > 0$ immediately implies that $\tau_B < 0$, i.e., bequests should be subsidized, the linear tax version of Farhi-Werning.

[22] They also will not directly respond to changes in bequest taxes leading to a smaller overall elasticity e_B.

[23] In our model, a future bequest tax change at date T has no impact on behavior until the first generation of donors (i.e., generation $T-1$) is hit.

assume inelastic earnings y_{Lti}, exogenous rate of return R, a constant and uniform discount factor δ, uniform felicity functions $u(c)$, and a utilitarian planner. Because labor supply is inelastic, we assume without loss of generality that $\tau_L = 0$ and that bequest taxes fund the demogrant E_t. The individual budget is $c_{ti} + b_{t+1i} = Rb_{ti}(1 - \tau_{Bt}) + \tau_{Bt}Rb_t$ with $b_{t+1i} \geq 0$. The individual first order condition implies that (regardless of corners where $b_{t+1i} = 0$),

$$u'(c_{ti})b_{t+1i} = \delta R(1 - \tau_{Bt+1})b_{t+1i}E_t u'(c_{t+1i}) \quad \text{and hence} \quad \bar{b}^{\text{left}}_{t+1} = \delta R(1 - \tau_{Bt+1})\bar{b}^{\text{received}}_{t+1} \tag{10}$$

where the second expression is obtained taking population averages and $\bar{b}^{\text{left}}_{t+1}$, $\bar{b}^{\text{received}}_{t+1}$ are again defined as the population averages of $u'(c_{ti})b_{t+1i}/[b_{t+1}Eu'(c_{ti})]$, $u'(c_{t+1i})b_{t+1i}/[b_{t+1}Eu'(c_{t+1i})]$.

In this model, the bequest tax redistributes from the wealthy to everybody which is desirable for insurance reasons (when $u'(c_{ti})$ is negatively correlated with b_{ti}) but is undesirable for efficiency reasons as τ_B affects wealth accumulation.

Optimum long-run τ_B for generation zero. Let $(\tau_{Bt})_{t\geq 0}$ be the sequence of tax rates that maximizes EV_0, i.e., expected utility of generation 0. We have:

$$EV_0 = \sum_{t\geq 0} \delta^t Eu(Rb_{ti}(1 - \tau_{Bt}) + \tau_{Bt}Rb_t + y_{Lti} - b_{t+1i}).$$

Assume that τ_{Bt} converges to τ_B. Consider a small reform $d\tau_B$ for all $t \geq T$ where T is large so that all variables have converged to their ergodic limit. As b_{ti} is chosen optimally, we can apply the envelope theorem, and the effect of the reform on EV_0 is:

$$dEV_0 = Rd\tau_B \sum_{t\geq T} \delta^t E[u'(c_{ti})(b_t - b_{ti})] + R\sum_{t\geq 0} \delta^t E[u'(c_{ti})]\tau_{Bt}db_t$$

The first term is the mechanical welfare effect (absent any behavioral response) while the second term reflects the welfare effect due to behavioral responses in savings behavior affecting tax revenue (and hence the demogrant). Importantly, note that the second sum starts at $t \geq 0$ as bequests may be affected before the reform takes place in anticipation. At the optimum,

$$0 = \frac{1}{R}\frac{dEV_0}{d\tau_B} = \sum_{t\geq T} \delta^t E[u'(c_{ti})(b_t - b_{ti})] + \sum_{t\geq 0} \delta^t E[u'(c_{ti})]b_t \frac{\tau_{Bt}}{1 - \tau_{Bt}} e_{Bt}, \tag{11}$$

with $e_{Bt} = \frac{1-\tau_{Bt}}{b_t}\frac{db_t}{d(1-\tau_B)}$ the elasticity of b_t with respect to the small reform $d\tau_B$ (for all $t \geq T$).

For $t \geq T$, τ_{Bt} changes by $d\tau_B$ and the savings decision is directly affected. When $t \to \infty$, e_{Bt} converges to the long-run elasticity e_B of b_t with respect to $1 - \tau_B$.[24] For $t < T$, τ_{Bt} does not

[24] This long-run elasticity e_B is calculated assuming that tax revenue is rebated lumpsum period by period.

change, hence savings decisions are only affected in anticipation of the future tax increase. In a model with no stochastic shocks (as in Chamley-Judd), the full path of consumption is shifted up for $t < T$ and then decreases faster for $t \geq T$. This implies that savings start responding from period 0 even for a very distant tax reform. In the stochastic model however, the anticipation response is attenuated as individuals hit the zero wealth constraint almost certainly as the horizon grows (see appendix A.1.4). Therefore, we can assume that e_{Bt} is non-zero only for t large at a point where τ_{Bt}, b_t and c_{ti} have converged to their long-run distribution. Hence, we can define the total elasticity e_B^T as the sum of the post-reform response elasticity e_B^{post} and the pre-reform anticipatory elasticity $e_B^{\text{anticip.}}$ as follows:

$$e_B^T = e_B^{\text{post}} + e_B^{\text{anticip.}} \quad \text{with } e_B^{\text{post}} = (1-\delta)\sum_{t \geq T} \delta^{t-T} e_{Bt} \quad \text{and } e_B^{\text{anticip.}} = (1-\delta)\sum_{t<T} \delta^{t-T} e_{Bt}. \tag{12}$$

e_B^T is the elasticity of the presented discounted value of the tax base with respect to a distant tax rate increase. e_B^{post} is the standard (discounted) average of the post-reform elasticities e_{Bt} while $e_B^{\text{anticip.}}$ is the sum of all the pre-reform behavioral elasticities e_{Bt}. We show in appendix A.1.4 that $e_B^{\text{anticip.}}$ becomes infinite when wage stochastic shocks disappear as in Chamley-Judd[25] but it is finite in the Aiyagari (1995) model with stochastic shocks. Naturally, $e_B^T \to e_B$ when $\delta \to 1$.[26] Using (10) and (12), we obtain from (11) the following optimal tax formulas.

Dynastic model long-run optimum, period 0 perspective, inelastic labor supply.

$$\tau_B = \frac{1 - \bar{b}^{\text{received}}}{1 - \bar{b}^{\text{received}} + e_B^T} \quad \textit{or equivalently} \quad \tau_B = \frac{1 - \frac{1}{\delta R}\bar{b}^{\text{left}}}{1 + e_B^T}. \tag{13}$$

where e_B^T, defined in (12), is the total (post-reform and anticipatory) elasticity of the present discounted value of aggregate bequests to a long-term distant pre-announced bequest tax increase, $\bar{b}^{\text{received}}$ and $\bar{b}^{\text{left}}$ are the population average of $b_{ti}u'(c_{ti})/[b_t E u'(c_{ti})]$ and $b_{t+1i}u'(c_{ti})/[b_{t+1} E u'(c_{ti})]$.

Six points are worth noting about formula (13). First, it shows that the standard equity-efficiency approach also applies to the dynastic model. The first expression in (13) takes the standard optimal linear tax rate form, decreasing in the elasticity e_B^T and decreasing with the distributional parameter $\bar{b}^{\text{received}}$. The key is simply to suitably define the elasticity e_B^T. As argued above, this elasticity is infinite in the Chamley-Judd model with no uncertainty so that our analysis nests the Chamley-Judd result. However, whenever the elasticity e_B^T is finite, the

[25] Importantly, in that case, $e_B^{\text{anticip.}}$ is infinite even in situations where the long-run elasticity e_B and hence e_B^{post} is finite as in the endogenous discount factor case of Judd (1985), Theorem 5, p. 79 (see appendix A.1.4).

[26] Numerical simulations could easily shed light on how $e_B^{\text{anticip.}}$, e_B^{post}, e_B change with the model specification and in particular the structure of stochastic shocks.

optimal tax rate is positive as long as $\bar{b}^{\text{received}} < 1$, i.e., bequests received are negatively correlated with marginal utility $u'(c_{ti})$ which is the expected case.[27]

Second, there is no double counting in the dynastic model from period 0 perspective. Hence, the cost of bequests taxation can be measured either on bequests receivers (first formula in (13)) or equivalently on bequests leavers (second formula in (13)). This shows that the optimal τ_B in the dynastic model takes the same form as (8), the long-run optimum with social discounting from Section 2, ignoring the welfare effect on bequests receivers, i.e., setting $\bar{b}^{\text{received}} = 0$.[28]

Third, with inelastic labor supply, the full optimum would be to set $\tau_L = 1$, i.e., provide perfect insurance, so that $\bar{b}^{\text{received}} = \bar{b}^{\text{left}} = 1$ and τ_B becomes useless.[29] The inelastic labor supply case is nevertheless pedagogically helpful.[30] As we show in appendix A.1.5, adding labor supply with heterogeneous felicity functions $u^{ti}(c, l)$ and considering a $d\tau_B, d\tau_L$ trade-off as in Section 2 modifies the optimal tax formula exactly as in our Section 2 analysis:

Dynastic model long-run optimum, period 0 perspective, and elastic labor supply.

$$\tau_B = \frac{1 - \frac{\bar{b}^{\text{received}}}{\bar{y}_L}\left[1 - \frac{e_L \tau_L}{1-\tau_L}\right]}{1 - \frac{\bar{b}^{\text{received}}}{\bar{y}_L}\left[1 - \frac{e_L \tau_L}{1-\tau_L}\right] + e_B^T} \quad \textit{or equivalently} \quad \tau_B = \frac{1 - \frac{1}{\delta R}\frac{\bar{b}^{\text{left}}}{\bar{y}_L}\left[1 - \frac{e_L \tau_L}{1-\tau_L}\right]}{1 + e_B^T}, \tag{14}$$

where e_L is an average elasticity of aggregate earnings with respect to $1 - \tau_L$ (see appendix A.1.5 for exact definition) and $\bar{y}_L$ is the population average of $y_{Lti} u'(c_{ti})/[y_{Lt} E u'(c_{ti})]$.

Fourth, optimal government debt management in the closed economy would deliver the Modified Golden rule $\delta R = 1$ and formulas (13) and (14) continue to hold (see appendix A.1.6). Hence, our model nests the Aiyagari (1995) analysis and provides a transparent economic intuition for why positive capital income taxation is desirable in his model.

Fifth, we can consider heterogeneous discount rates δ_{ti}. Formula (13) still applies with $\bar{b}^{\text{received}} = \lim_T \frac{\sum_{t \geq T} E[\delta_{1i} \cdot\ldots\cdot \delta_{ti} u'(c_{ti}) b_{ti}]}{\sum_{t \geq T} E[\delta_{1i} \cdot\ldots\cdot \delta_{ti} u'(c_{ti})] b_t}$. Hence, $\bar{b}^{\text{received}}$ puts weight on consistently altruistic dynasties, precisely those that accumulate wealth so that $\bar{b}^{\text{received}} > 1$ and $\tau_B < 0$ is likely. In that case, the period 0 criterion puts no weight on individuals who had non-altruistic ancestors. This fits with aristocratic values, but is the polar opposite of modern meritocratic values.

[27] This point on the sign of optimal long-run bequest taxation was made by Chamley (2001) although he did not derive an optimal tax formula. He also crafts an example showing that $\bar{b}^{\text{received}} > 1$ is theoretically possible.

[28] Naturally, $\tau_L = e_L = 0$ here. Note also that $\bar{y}_L$ is replaced by one because the trade-off here is between the bequest tax and the demogrant (instead of the labor tax as in Section 2).

[29] With no uncertainty, the steady-state requires $\delta R = 1$ so that the second formula (13) also delivers $\tau_B = 0$.

[30] As discussed below, our results easily extend to heterogeneous discount factors δ_{ti}. In that case, even with no earnings heterogeneity (or $\tau_L = 1$), inequality arises due to bequests and the optimal τ_B is non-zero.

Sixth, adding Pareto weights ω_{0i} that depend on initial position delivers exactly the same formula as the long-run position of each individual is independent of the initial situation. This severely limits the scope of social welfare criteria in the period 0 perspective model.

Optimum long-run τ_B in the steady-state. Finally, to come back full circle to our initial analysis in Section 2, it is instructive to consider the optimum long-run τ_B that maximizes the steady-state utility of the dynastic model (instead of period 0 utility). We consider elastic labor supply as in Section 2 (without adding any notational complication). The government solves

$$\max_{\tau_L,\tau_B} EV_\infty = \sum_{t\geq 0} \delta^t E[u^{ti}(Rb_{ti}(1-\tau_B) + w_{ti}l_{ti}(1-\tau_L) + E - b_{t+1i}, l_{ti})],$$

where we assume (w.l.o.g) that the steady-state has been reached in period 0. b_{0i} is given to the individual (but depends on τ_B) while b_{ti} for $t \geq 1$ (and l_{ti} for $t \geq 0$) are chosen optimally so that the envelope theorem applies. Therefore, first order condition with respect to τ_B is:

$$0 = R(1-\tau_B)E\left[u_c^{ti}(c_{0i}, l_{i0})db_{0i}\right] + \sum_{t\geq 0} \delta^t RE\left[u_c^{ti}(c_{ti}, l_{ti}) \cdot (-b_{ti}d\tau_B + y_{Lti}d\tau_L - \tau_L dy_{Lti})\right]. \quad (15)$$

Equation (4) linking $d\tau_L$ and $d\tau_B$ continues to hold in this dynastic model with long-run steady state elasticities e_B and e_L. The anticipatory elasticity disappears in the steady-state model and e_B is the long-run elasticity as in Section 2. Hence, the only difference with Section 2 is that, except for the initial bequest effect db_{0i}, all terms are repeated (with discount factor δ). Hence, this is equivalent to discounting the initial bequest effect by a factor $1-\delta$ so that:

Dynastic model long-run optimum, steady-state perspective.

$$\tau_B = \frac{1 - \left[1 - \frac{e_L\tau_L}{1-\tau_L}\right] \cdot \left[\frac{(1-\delta)\bar{b}^{\text{received}}}{\bar{y}_L}(1+\hat{e}_B) + \frac{1}{R}\frac{\bar{b}^{\text{left}}}{\bar{y}_L}\right]}{1 + e_B - \left[1 - \frac{e_L\tau_L}{1-\tau_L}\right]\frac{(1-\delta)\bar{b}^{\text{received}}}{\bar{y}_L}(1+\hat{e}_B)}. \quad (16)$$

Four points are worth noting. First, the formula is very close to (7), the only difference being the discount factor $1-\delta$ for $\bar{b}^{\text{received}}$. This factor appears because all the other effects are magnified by a factor $\frac{1}{1-\delta} = 1 + \delta + \delta^2 + ...$ Hence, for given elasticities and distributional parameters, the dynastic case actually makes the optimal τ_B *larger* than our initial model.[31]

Second, if stochastic shocks vanish, then $e_B = \infty$ and $\tau_B = 0$. This nests the steady-state maximization version of Chamley and Judd (presented in Piketty, 2000 p. 444) that

[31]The steady-state also makes the optimal τ_B *smaller* than period 0 formula by adding concerns for bequest receivers. Farhi and Werning (2010) obtain a similar result in a NDPF model (see appendix A.4).

naturally delivers a zero optimum when the supply elasticity of capital is infinite.[32] With stochastic shocks, the elasticity e_B is finite. For given utility functional forms and stochastic wage processes, calibrations of the dynastic stochastic model are likely to deliver larger e_B than our bequests in the utility Section 2 model, likely yielding lower τ_B. The power of our sufficient statistics approach is to show precisely the key relevant parameter. As long as e_B is empirically known, the primitives of the model (dynastic vs. bequest loving) are largely irrelevant.

Third, the steady-state case allows to consider general welfare weights ω_{0i}, a valuable flexibility. In (15), the sums over t are no longer identical terms as the correlation of social marginal welfare weights $\omega_{i0}u_c(c_{ti}, l_{ti})$ with b_{ti} and y_{Lti} changes with t. Hence, in that case $\frac{1}{1-\delta}$, $\bar{b}^{\text{left}}$, and $\bar{y}_L$ have to be replaced by with $\frac{1}{1-\delta} = \sum_{t\geq 0}\delta^t \frac{E[\omega_{0i}u_c(c_{ti},l_{ti})]}{E[\omega_{0i}u_c(c_{0i},l_{0i})]}$, $\bar{b}^{\text{left}} = \frac{\sum_{t\geq 0}\delta^t E[\omega_{0i}u_c(c_{ti},l_{ti})b_{ti}]}{\sum_{t\geq 0}\delta^t E[\omega_{0i}u_c(c_{0i},l_{0i})]b_t}$, $\bar{y}_L = \frac{\sum_{t\geq 0}\delta^t E[\omega_{0i}u_c(c_{ti},l_{ti})y_{Lti}]}{\sum_{t\geq 0}\delta^t E[\omega_{0i}u_c(c_{0i},l_{0i})]y_{Lt}}$. In the zero-receiver Rawlsian optimum, all the terms in $\bar{b}^{\text{received}}$ vanish so that (7) applies. The term $\bar{b}^{\text{left}}$ not only measures the correlation of social marginal utility with bequests left among zero-receivers but also among descendants of zero receivers.[33]

Finally, it is possible to write a fully general model $V^{ti} = u^{ti}(c, \underline{b}, b, l) + \delta_{ti}V^{t+1i}$ that encompasses many possible bequests motivations. The optimal formula in the steady-state continues to take the same general shape we have presented, although notations are more cumbersome.

4 Numerical Calibrations

We use wealth surveys for France (Enquête Patrimoine 2010) and the U.S. (Survey of Consumer Finances 2010) to calibrate the general steady-state formula (see appendix A.5 for details)

$$\tau_B = \frac{1 - \left[1 - e_L\frac{\tau_L}{1-\tau_L}\right] \cdot \left[\frac{\bar{b}^{\text{received}}}{\bar{y}_L}(1+\hat{e}_B) + \frac{\nu}{R/G}\frac{\bar{b}^{\text{left}}}{\bar{y}_L}\right]}{1 + e_B - \left[1 - e_L\frac{\tau_L}{1-\tau_L}\right]\frac{\bar{b}^{\text{received}}}{\bar{y}_L}(1+\hat{e}_B)}, \tag{17}$$

which incorporates growth (G) and wealth loving motivations (ν). We consider the following benchmark values for the parameters: $e_B = \hat{e}_B = 0.2$, $e_L = 0.2$, $\tau_L = 30\%$, $\nu = 70\%$, $R/G = e^{(r-g)H} = 1.82$ with $r - g = 2\%$ and $H = 30$ years. The distributional parameters $\bar{b}^{\text{received}}$, $\bar{b}^{\text{left}}$ and $\bar{y}_L$ are computed within the population of individuals aged 70 and above, and for each percentile of the distribution of bequest received.[34] We use retrospective questions about bequest

[32] The finite long-run elasticity case presented by Judd (1985), Section 5 would deliver a positive optimal τ_B maximizing steady-state utility even though it delivers a zero optimal long-run τ_B from period 0 perspective as the anticipatory elasticity component remains infinite as discussed above.

[33] As generations mean revert, $\bar{b}^{\text{left}}$ is in-between the zero-receivers case and the limit utilitarian case. If δ is relatively small (e.g., $\delta < 1/2$), the quantitative difference is likely small.

[34] We focus on older cohorts because they have already received bequests from their parents, and will soon leave bequests to their children. Hence, we can estimate both the distribution of bequests received and left.

and gift receipts available in both surveys to compute $\bar{b}^{\text{received}}$. We use questions about current net wealth to estimate $\bar{b}^{\text{left}}$. We use the information available on the sum of wage income, self-employment income, and pension income (usually proportional to past labor income) to compute $\bar{y}_L$.[35] All distributional parameters are estimated at the individual level.[36]

The main results are summarized on Figure 1, where we report optimal linear inheritance tax rates by percentile of the distribution of bequest received. We find that in both countries the optimal tax rate is about 50%-60% for the bottom 70% of the population, then falls abruptly and becomes negative within the top 20% of inheritors (particularly for the top 10%).[37] Although our results should be viewed as exploratory, a number of conclusions appear to be robust.

Most importantly, bottom 70% bequest receivers have virtually the same most preferred bequest tax rate because of the very large concentration of inherited wealth. In every country for which we have data, the bottom 50% share in aggregate inherited wealth is typically about 5% or less, and the top 10% share is 60%-70% or more (Piketty, 2011, p.1076). Hence, $\bar{b}^{\text{received}}$ is very close to 0% for the bottom 50% of the distribution of received bequests, and barely higher for the next 20%. In both countries, bottom 50% bequest receivers have average labor incomes fairly close to national averages (with $\bar{y}_L$ around 90%), but leave substantially less wealth than average to their children (with $\bar{b}^{\text{left}}$ around 60%-70%). Hence, in both countries the optimal inheritance tax rate is fairly large for bottom bequest receivers: even though they enjoy leaving bequests, it is in their interest to tax bequests at relatively large rates, so to as reduce their labor tax burden and accumulate more wealth. Optimal tax rates are close in both countries - except that they start falling around percentile 70 in France and percentile 80 in the US.[38]

Our results illustrate the fact that inheritance taxation involves deeply conflicting economic interests: bottom groups benefit from high inheritance tax rates, but relatively large groups at the top would benefit from inheritance subsidies. They also show the importance of beliefs, ideology, and political discourses regarding wealth mobility: the optimal tax rates reported on

[35]We repeated the same computations separately for individuals aged 60-69, 70-79, and 80-89, and found almost identical results. Using transmissible net wealth (excluding pension funds) rather than net wealth or using information on past occupation to estimate $\bar{y}_L$ had very small effects on estimates.

[36]Wealth of married individuals is defined as household wealth divided by two. Similarly, bequest received is defined as the sum of all bequests and gifts received by both spouses divided by two.

[37]We put a lower bound $\tau_B = -20\%$ for readability as the optimum is infinitely negative in upper percentiles.

[38]This is due to the larger concentration of inherited wealth in the U.S. (i.e. $\bar{b}^{\text{received}}$ remains very close to 0% until percentile 80 in the U.S., while it becomes significant after percentile 70 in France). Conversely, $\bar{b}^{\text{left}}$ among bottom 50% receivers is larger in the U.S., suggesting higher wealth mobility. Some of these differences between the two countries might reflect reporting biases (namely, bequests received might be particularly under-reported in the U.S., which would explain both findings) and should be further analyzed in future research. See Appendix A.5 for a discussion.

Figure 1 correspond to observed mobility patterns and would be much lower if a significant part of the population was over-optimistic about their prospects for leaving large bequests.

Table 1 considers alternative parameters around our basic specification (see appendix A.5).

First, regarding the bequest elasticity e_B, we choose a benchmark value $e_B = 0.2$. Using U.S. time and cross-section variations, Kopczuk and Slemrod (2001) find elasticities e_B around $0.1 - 0.2$.[39] There remains considerable uncertainty about e_B. With $e_B = 0$, the optimal inheritance tax rate for bottom receivers would be about 70% (rather than 60%). With $e_B = 0.5$, it would be about 50%. Even with an elasticity $e_B = 1$, which seems implausibly high, the optimal inheritance tax rate would still be about 35% in both countries (Figures A1-A2).

Second, regarding the capitalization factor R/G, we choose a benchmark value $R/G = e^{(r-g)H} = 1.82$, which corresponds to $r - g = 2\%$ and $H = 30$ years. Historically, the difference between the average rate of return to wealth and the growth rate has been closer to 3%-4% or even higher (Piketty 2011, Table II, p.1122). With $r - g = 3\%$, optimal inheritance tax rates would be close to 70%, both in France and in the US. Conversely, assuming $r - g = 0\%$, i.e., $R/G = 1$, which can be interpreted as the case with dynamic efficiency and optimal capital accumulation, optimal inheritance tax rates fall to about 40% in both countries (Figures A3-A4).

Third, regarding the strength of the bequest motive, we use a benchmark value $\nu = 70\%$. According to Kopczuk and Luton (2005), there is substantial heterogeneity in the distribution of motives for wealth accumulation. The average fraction of the population with a bequest motive is between one half and two thirds, hence $\nu = 70\%$ is on the high end. With $\nu = 0\%$, i.e. in the complete absence of bequest motives, e_B is the sole limiting factor for optimal tax rates, which would then be over 80%. Conversely, with $\nu = 100\%$, i.e. if old age wealth accumulation is entirely driven by bequest motives, the optimal inheritance rate would fall slightly below 50% (Figures A5-A6).

Fourth, optimal bequest tax rates are only weakly increasing with the labor elasticity e_L.

Fifth, to illuminate the crucial role played by wealth inequality and mobility, we also provide in Appendix A.5 estimates using the micro files of estate tax returns collected by Piketty, Postel-Vinay, and Rosenthal (2011) in the Paris archives over the 1872-1937 period, a time characterized by large inheritance flows and extreme wealth concentration (with over 90% of aggregate inheritance received by top 10% successors). This is highly reliable, exhaustive ad-

[39] Preliminary computations using time and cross section variations in French inheritance tax rates (e.g. in the French system childless individuals pay a lot more bequest taxes than individuals with children) also suggest that e_B is relatively small (at most $e_B = 0.1 - 0.2$).

ministrative data covering wealth over two generations. We find that $\bar{b}^{\text{left}}$ is as low as 20%-30% for the bottom 80% receivers, implying very high optimal inheritance tax rates - typically above 80% for the benchmark values parameters used here.

Sixth, it is easy to extend the optimal linear tax formula to nonlinear bequest taxation that takes the form of a simple two-bracket tax with a linear tax rate above an exemption threshold, a reasonable approximation to actual schedules. Our formula carries over virtually unchanged by replacing bequests by taxable bequests above the exemption threshold in our formulas (see appendix A.6.1).[40] Figure 2 shows that in both countries the optimal top tax rate above an exemption level (of \$500,000 or \$1m) is higher than the optimal linear inheritance tax rate, particularly in France where bottom 50% bequest receivers have a relatively small probability to leave bequests above such levels.[41] It is worth noting that these high top inheritance tax rates - say around 60%-70% - are very much in line with historical experience, especially in Anglo-Saxon countries from the 1930s to the 1980s, when top estate tax rates were systematically above 60% (Figure 3). The decline of US top rates since the 1980s could be due to a shift in political power away from the bottom 80% and toward the top 10%. Finally, Figure 2 shows that a smaller minority at the top opposes top bequests taxes than linear bequest taxes, explaining perhaps why actual bequest taxes often have fairly large exemption levels.

5 Conclusion and Extensions

This paper has derived robust optimal inheritance tax formulas expressed in terms of sufficient statistics. This approach casts fruitful light on the problem and unifies previous seemingly disparate results. Our punchline is that, in accordance to the public debate, the optimal tax rate trades-off equity and efficiency and this trade-off is non-degenerate if the elasticity of bequests with respect to taxation is not infinite and inheritances matter for life-time resources and social preferences. If the concentration of inheritances is high, the elasticity is low, and the government favors those with little inheritance, the optimal tax rate will be high. Our analysis could be extended in a number of directions.

First, solving the full non-linear optimum (instead of only the two-bracket case) would be valuable. This complicates the analysis but does not radically change the optimal tax problem.

[40]It is computationally more difficult to solve for the optimal threshold (and even more so for the optimal many-bracket nonlinear tax schedule). Hence, we take the exemption threshold as given.

[41]This could again be partly due to reporting biases (Appendix A.5).

Second, if the government can use debt, labor taxation τ_L is exactly equivalent to a consumption tax τ_C even in the presence of bequests provided the government compensates individuals for initial wealth implicitly taxed when switching from labor to consumption tax (appendix A.6.2). Hence, the same formulas for τ_B apply when considering the trade-off between bequest taxation and consumption taxation (instead of labor taxation). The view that consumption taxation can successfully tax wealthy idle heirs is illusory because, with labor income taxation, those wealthy heirs would have received smaller inheritances to start with. With nonlinear taxation, the full equivalence between labor and consumption tax naturally breaks down. But it is still the case that consumption taxation is a poor instrument to target inheritors, unless of course one assumes that inheritance taxes are not available.[42]

Third, we have limited ourselves to the analysis of capitalized inheritance taxation. That is, the same tax rate τ_B is used to tax bequest b_i and lifetime return to bequest $(R-1)b_i$. In our one-period model, a capitalized inheritance tax τ_B is actually equivalent to a pure capital income tax τ_K if $R(1-\tau_B) = 1+(R-1)(1-\tau_K)$ so that our results can also be interpreted as a theory of capital income taxation. In practice, capital income and wealth taxation is much more significant than bequest taxation. Capital income taxation raises other interesting issues. Firstly, as we have seen, life-cycle savings taxation distorts inter-temporal choices with no redistributive benefits. This would push toward taxing solely bequests and not tax at all capital income. Secondly however, if there is a fuzzy frontier between capital income and labor income, zero capital income taxation would lead to re-characterization of labor income into capital income. To close this loophole, the government can set $\tau_K = \tau_L$ and then decrease τ_B so that the total tax wedge on capitalized bequests remains the same as in our formulas (see Piketty and Saez, 2012). Thirdly, there might be other reasons why capital income taxation could be desirable. Bequests taxation might force inefficient sale of indivisible assets in the presence of credit constraints (or might be more disliked than annual lower capital income or wealth taxes due to fiscal illusion). More importantly, rates of return on capital vary widely across individuals. To the extent that such risk is not optimally diversified, capital income taxation could be desirable for rate of return insurance reasons. That is, with capital market imperfections, lifetime capital income and wealth taxation might be the efficient way to implement optimal inheritance taxes (Piketty and Saez 2012 present a basic model along those lines).

[42] This simple point (i.e. with ill functionning inheritance and capital taxes one can use progressive consumption taxes to tax wealthy successors) was first made by Kaldor (1955). See Piketty and Saez (2012, Appendix B.4).

References

Aiyagari, S. Rao 1995. "Optimal Capital Income Taxation with Incomplete Markets, Borrowing Constraints, and Constant Discounting," *Journal of Political Economy* 103(6), 1158-75.
Atkinson, Anthony and Agnar Sandmo 1980. "Welfare Implications of the Taxation of Savings," *Economic Journal*, 90, 529-549.
Atkinson, Anthony and Joseph E. Stiglitz 1976. "The Design of Tax Structure: Direct Versus Indirect Taxation," *Journal of Public Economics* 6(1–2): 55–75.
Chamley, Christophe 1986. "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives," *Econometrica* 54(3), 607–22.
Chamley, Christophe 2001. "Capital Income Taxation, Wealth Distribution and Borrowing Constraints," *Journal of Public Economics* 79, 55–69.
Deaton, Angus 1979. "Optimally Uniform Commodity Taxes." *Economic Letters*, 2, 357–361.
Diamond, Peter and James A. Mirrlees 1971. "Optimal Taxation and Public Production: Production Efficiency." *American Economic Review,* 61(1), 8–27.
Diamond, Peter and Ivan Werning 2013. "Dynamic Taxation", *Handbook of Public Economics*, Volume 5, Elsevier.
Farhi, Emmanuel and Ivan Werning 2010. "Progressive Estate Taxation," *Quarterly Journal of Economics*, 125(2), 635-673.
Fleurbaey, Marc 2008. *Fairness, Responsability and Welfare*, Oxford University Press.
Judd, Kenneth 1985. "Redistributive Taxation in a Simple Perfect Foresight Model." *Journal of Public Economics*, 28(1), 59–83.
Kaldor, Nicholas 1955. *An Expenditure Tax*, Unwin University Books, 248p.
Kaplow, Louis 2001. "A Framework for Assessing Estate and Gift Taxation", in Gale, W., J. Hines, J. Slemrod (eds.) *Rethinking estate and gift taxation* D.C.: Brookings Institution Press.
King, Mervyn 1980. "Savings and Taxation", in G. Hughes and G. Heal, eds., *Public Policy and the Tax System* (London: George Allen Unwin), 1-36.
Kocherlakota, Narayana 2010. *The New Dynamic Public Finance* (Princeton: Princeton University Press).
Kopczuk, Wojciech 2013. "Taxation of Transfers and Wealth," *Handbook of Public Economics*, Volume 5, Elsevier.
Kopczuk, Wojciech and Joseph Lupton 2007. "To Leave or Not to Leave: The Distribution of Bequest Motives", *Review of Economic Studies*, 74(1), 207-235.
Kopczuk, Wojciech and Joel Slemrod 2001. "The Impact of the Estate Tax on Wealth Accumulation and Avoidance Behavior", in *Rethinking Estate and Gift Taxation*, W. Gale, J.R. Hines and J. Slemrod eds., The Brookings Institution, 299–349.
Piketty, Thomas 2000. "Theories of Persistent Inequality and Intergenerational Mobility", In Atkinson A.B., Bourguignon F., eds. *Handbook of Income Distribution*, (North-Holland).
Piketty, Thomas 2011. "On the Long-Run Evolution of Inheritance: France 1820-2050", *Quarterly Journal of Economics*, 126(3), 1071–1131.
Piketty, Thomas, Gilles Postel-Vinay and Jean-Laurent Rosenthal, 2011. "Inherited vs Self-Made Wealth: Theory and Evidence from a Rentier Sociery (1872-1937)", PSE WP.
Piketty, Thomas and Emmanuel Saez 2012 "A Theory of Optimal Capital Taxation," NBER Working Paper No. 17989.
Piketty, Thomas and Emmanuel Saez 2013. "Optimal Labor Income Taxation", *Handbook of Public Economics*, Volume 5, Elsevier.
Piketty, Thomas and Gabriel Zucman 2012. "Capital is Back: Wealth-Income Ratios in Rich Countries, 1870-2010,"PSE Working Paper.
Saez, Emmanuel 2002. "The Desirability of Commodity Taxation under Nonlinear Income Taxation and Heterogeneous Tastes," *Journal of Public Economics*, 83, 217-230.

Table 1: Optimal inheritance tax rate τ_B calibrations

	Elasticity e_B=0 (low-end estimate)		Elasticity e_B=0.2 (middle-end estimate)		Elasticity e_B=0.5 (high-end estimate)		Elasticity e_B=1 (extreme estimate)	
	France	US	France	US	France	US	France	US
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
0. Basic Specification: optimal tax for bottom 50% receivers, r-g=2% (R/G=1.82), ν=70%, e_L=0.2, no exemption (linear tax τ_B)								
P0-50, r-g=2%, ν=70%, e_L=.2	76%	70%	63%	59%	50%	47%	38%	35%
1. Optimal tax rate for other groups by percentile of bequests received								
P50-70	75%	70%	62%	59%	48%	47%	35%	35%
P70-90	45%	60%	31%	46%	16%	31%	2%	17%
P90-95	-283%	-43%	-330%	-84%	-376%	-126%	-423%	-167%
2. Sensitivity to capitalization factor R/G=$e^{(r-g)H}$								
r-g=0% (R/G=1)	56%	46%	46%	38%	37%	31%	28%	23%
r-g=3% (R/G=2.46)	82%	78%	68%	65%	55%	52%	41%	39%
3. Sensitivity to bequests motives ν								
ν=1 (100% bequest motives)	65%	58%	54%	48%	43%	39%	33%	29%
ν=0 (no bequest motives)	100%	100%	83%	83%	67%	67%	50%	50%
4. Sensitivity to labor income elasticity e_L								
e_L=0	73%	68%	61%	56%	49%	45%	37%	34%
e_L=0.5	79%	75%	66%	62%	53%	50%	40%	37%
5. Optimal top tax rate above positive exemption amount								
Exemption amount: 500,000	88%	73%	65%	58%	46%	44%	32%	31%
Exemption amount: 1,000,000	92%	73%	66%	57%	46%	43%	30%	31%

This table presents simulations of the optimal inheritance tax rate τ_B using formula (17) from the main text for France and the United States and various parameter values. In formula (17), we use τ_L=30% (labor income tax rate). Parameters $b^{received}$ b^{left}, y_L are obtained from the survey data (see appendix A.5).

Figure 1: Optimal linear inheritance tax rates, by percentile of bequest received (calibration of optimal tax formulas using 2010 micro data)

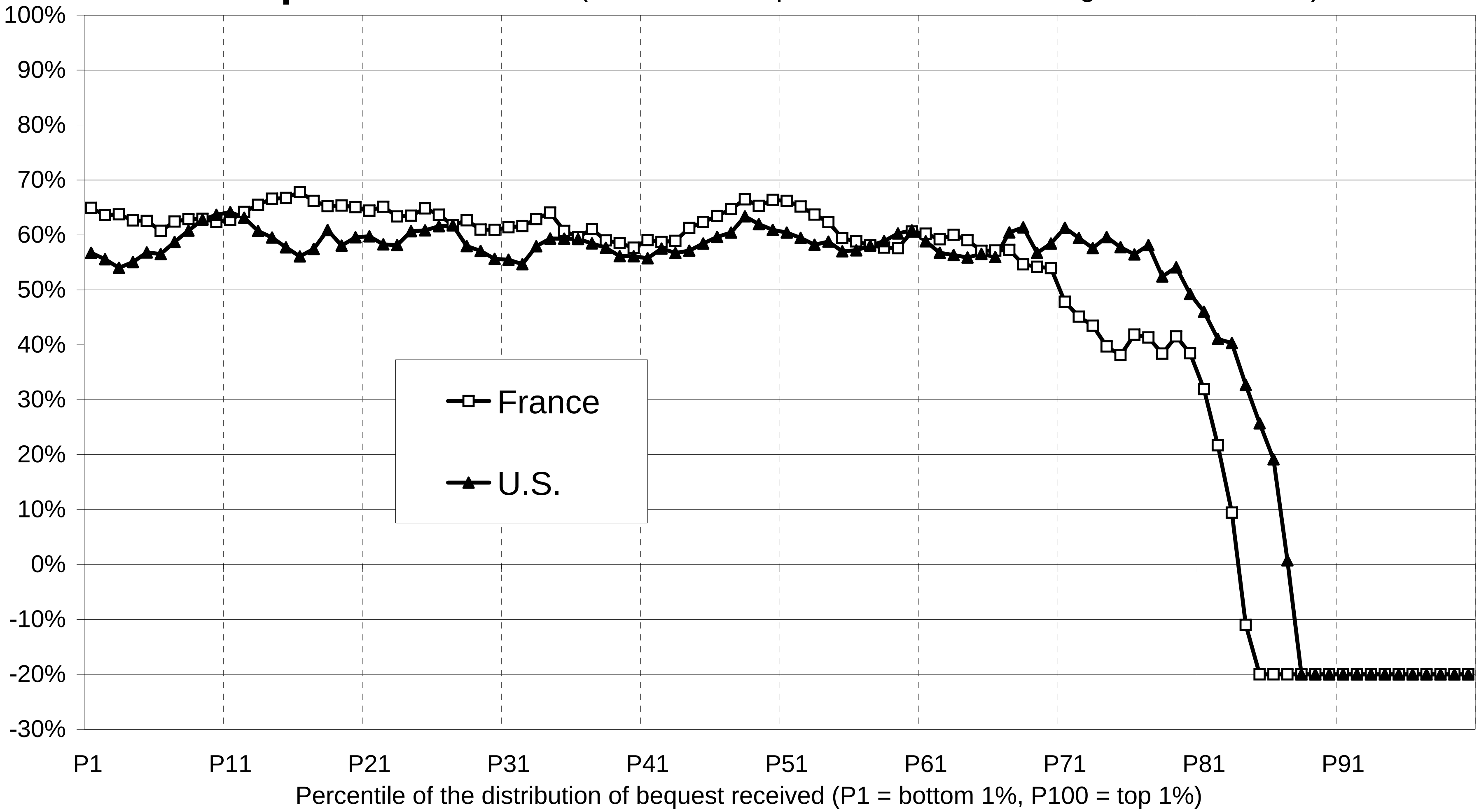

Figure 2: Optimal top inheritance tax rates, by percentile of bequest received (1m€ or $+) (calibration using 2010 micro data)

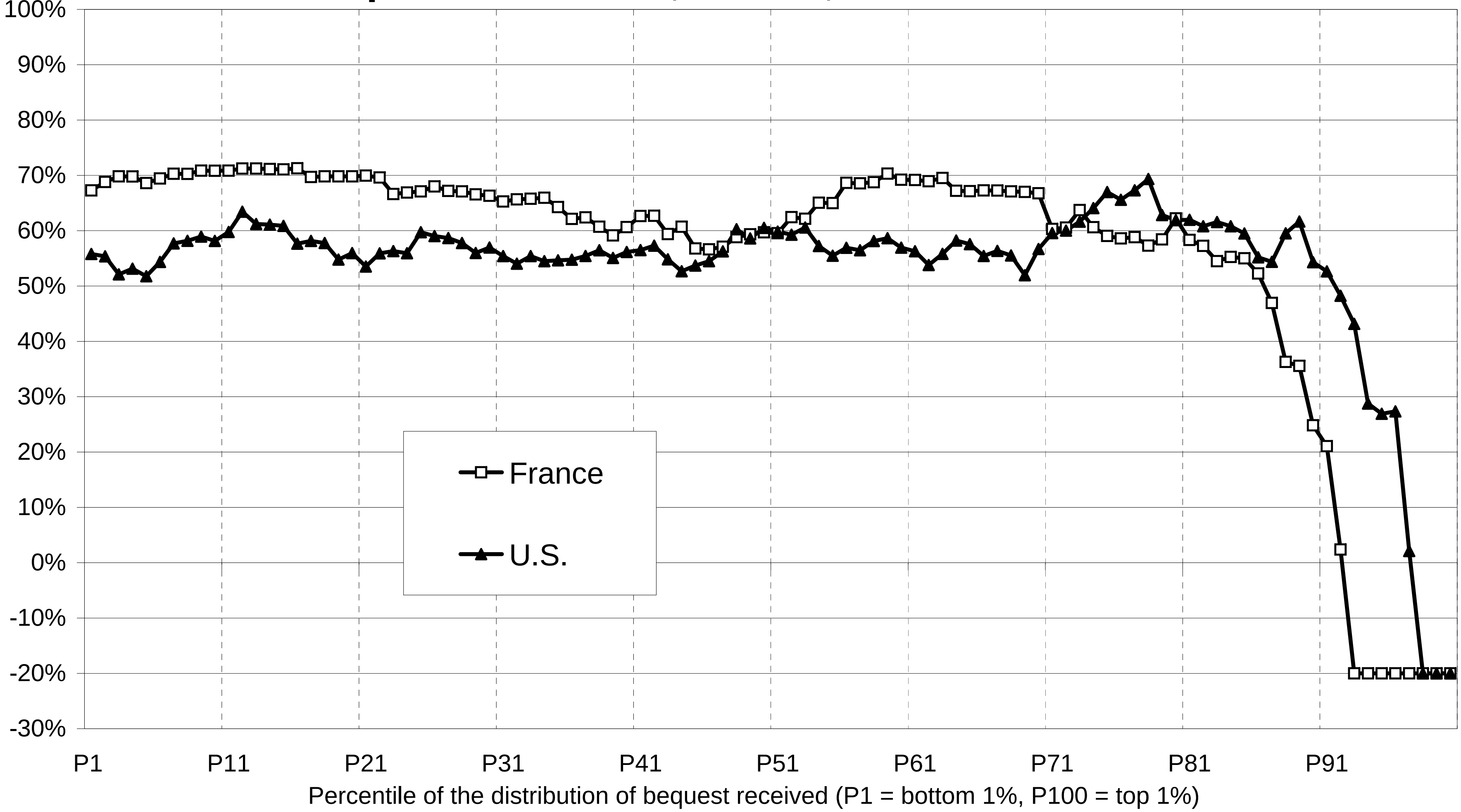

Figure 3: Observed top inheritance tax rates 1900-2011

Web Appendix of
A Theory of Optimal Inheritance Taxation

Thomas Piketty and Emmanuel Saez

A.1 Technical Details and Omitted Proofs from the Main Text

A.1.1 Formula (8) with Social Discounting Δ

We consider the small open economy with exogenous R, period-by-period budget balance, and social discounting at rate Δ. The government maximizes

$$SWF = \sum_{t \geq 0} \Delta^t \int_i \omega_{ti} V^{ti}(Rb_{ti}(1-\tau_{Bt}) + w_{ti}l_{ti}(1-\tau_{Lt}) + E_t - b_{t+1i}, Rb_{t+1i}(1-\tau_{Bt+1}), l_{ti}),$$

subject to period-by-period budget balance $E_t = \tau_{Bt} Rb_t + \tau_{Lt} y_{Lt}$. Consider again a reform $d\tau_B$ so that $d\tau_{Bt} = d\tau_B$ for all $t \geq T$ (and correspondingly $d\tau_{Lt}$ to maintain budget balance and keeping E_t constant). We assume that T is large enough so that all variables have converged for $t \geq T$. As b_{t+1i}, l_{ti} maximize individual utility, using the envelope theorem, the effect of the tax reform on social welfare is

$$dSWF = \sum_{t \geq T} \Delta^t \int_i \omega_{ti} V_c^{ti} \cdot (Rdb_{ti}(1-\tau_{Bt}) - Rb_{ti} d\tau_B - d\tau_{Lt} y_{Lti}) + \sum_{t \geq T-1} \Delta^t \int_i \omega_{ti} V_{\underline{b}}^{ti} \cdot (-d\tau_B Rb_{t+1i}).$$

To rewrite this equation in terms of elasticities of b_t and y_{Lt} with respect to $1-\tau_B$ and $1-\tau_L$, we can define e_{Bt} as the elastic response of b_t to the tax reform $d\tau = (d\tau_{Bt} = d\tau_B, d\tau_{Lt})_{t \geq T}$, so that $\frac{db_t}{b_t} = -e_{Bt} \frac{d\tau_B}{1-\tau_B}$ where db_t is the aggregate bequest response to the full reform $d\tau$. Note that the response of b_t starts only in period T (as bequest leavers care only about the net-of-tax bequests they leave). The response builds over generations and eventually converges to the long-run e_B, as defined in (3). We similarly define the elasticity e_{Lt} so that $\frac{dy_{Lt}}{y_{Lt}} = -e_{Lt} \frac{d\tau_{Lt}}{1-\tau_L}$ where dy_{Lt} is the labor supply response to the full reform $d\tau$. Period-by-period budget balance requires:

$$Rb_t d\tau_B \left(1 - e_{Bt} \frac{\tau_B}{1-\tau_B}\right) = -d\tau_{Lt} y_{Lt} \left(1 - e_{Lt} \frac{\tau_L}{1-\tau_L}\right). \qquad \text{(A1)}$$

Using the individual first order condition $V_c^{ti} = R(1-\tau_B) V_{\underline{b}}^{ti}$ when $b_{t+1i} > 0$, along with the budget balance equation (A1) allows to rewrite the first order condition $dSWF = 0$ as

$$0 = \sum_{t \geq T} \Delta^t \int_i g_{ti} \left[-d\tau_B Rb_{ti}(1+e_{Bti}) + \frac{1 - e_{Bt}\tau_B/(1-\tau_B)}{1 - e_{Lt}\tau_L/(1-\tau_L)} \frac{y_{Lti}}{y_{Lt}} Rb_t d\tau_B \right] - \sum_{t \geq T-1} \Delta^t \int_i g_{ti} d\tau_B \frac{b_{t+1i}}{1-\tau_B},$$

The first term captures the negative effect of $d\tau_B$ on bequest received (the direct effect and the dynamic effect via reduced pre-tax bequests), the second term captures the positive effect through reduced labor income tax, the third term captures the negative effect on bequest leavers.

Importantly, note that the third term is a sum starting at $T-1$ (instead of T), as the reform hurts bequest leavers starting in generation $T-1$.

As everything has converged for $t \geq T$, dividing by $Rb_t d\tau_B$ and denoting by $\bar{y}_L$, $\bar{b}^{\text{received}}$, $\bar{b}^{\text{left}}$, the population averages of $g_{ti} \cdot y_{Lti}/y_{Lt}$, $g_{ti} \cdot b_{ti}/b_t$, $g_{ti} \cdot b_{t+1i}/b_t$, and $\hat{e}_{Bt} = \int_i g_{ti} b_{ti} e_{Bti} / \int_i g_{ti} b_{ti}$, the first order condition is rewritten as:

$$0 = -\sum_{t \geq T} \Delta^t \bar{b}^{\text{received}}(1+\hat{e}_{Bt}) + \sum_{t \geq T} \Delta^t \frac{1 - e_{Bt}\tau_B/(1-\tau_B)}{1 - e_{Lt}\tau_L/(1-\tau_L)} \bar{y}_L - \sum_{t \geq T-1} \Delta^t \frac{\bar{b}^{\text{left}}}{R(1-\tau_B)},$$

We define $e_B = (1-\Delta)\sum_{t \geq T} \Delta^{t-T} e_{Bt}$, $\hat{e}_B = (1-\Delta)\sum_{t \geq T} \Delta^{t-T} \hat{e}_{Bt}$ as the discounted average of the e_{Bt} and $\hat{e}_{Bt}$. We then define e_L so that:

$$\frac{1 - e_B\tau_B/(1-\tau_B)}{1 - e_L\tau_L/(1-\tau_L)} = (1-\Delta)\sum_{t \geq T} \Delta^{t-T} \frac{1 - e_{Bt}\tau_B/(1-\tau_B)}{1 - e_{Lt}\tau_L/(1-\tau_L)}.$$

Naturally, in the case e_{Lt} constant in t, then we have $e_{Lt} \equiv e_L$. Such would be the case with iso-elastic quasi-linear utility functions of the form $V^{ti}(c, \underline{b}, l) = U^{ti}(c - l^{1+1/e_L}, \underline{b})$ where labor supply depends solely on the net-of-tax wage rate and has no income effects. In the general case with income effects, e_{Lt} might vary slightly with t as the building response db_{ti} creates income effects varying with t. Using those definitions we can rewrite the first order condition as:

$$0 = -\bar{b}^{\text{received}}(1+\hat{e}_B) + \frac{1 - e_B\tau_B/(1-\tau_B)}{1 - e_L\tau_L/(1-\tau_L)} \bar{y}_L - \frac{\bar{b}^{\text{left}}}{\Delta R(1-\tau_B)},$$

where the Δ in the denominator of the third term appears because the sum for the third term starts at $T-1$ instead of T. Re-arranging this expression leads immediately to formula (8). Finally, note that the elasticities used here converge to the long-run elasticities from Section 2 when $\Delta = 1$, i.e., when there is no social discounting. Q.E.D.

A.1.2 Case with Endogenous Factor Prices

Individual ti solves the problem

$$\max_{c_{ti}, b_{t+1i} \geq 0} V^{ti}(c_{ti}, \underline{R}_t b_{t+1i}, l_{ti}) \text{ s.t. } c_{ti} + b_{t+1i} = \underline{R}_t b_{ti} + \underline{w}_t v_{ti} l_{ti} + E_t. \qquad \text{(A2)}$$

where $\underline{R}_t = R_t(1-\tau_{Bt})$ and $\underline{w}_t = w_t(1-\tau_{Lt})$ are the after-tax factor prices, v_{ti} is the work ability of individual ti so that her pre-tax wage is $w_t v_{ti}$. The individual first order condition is $V_c^{ti} = \underline{R}_{t+1} V_{\underline{b}}^{ti}$ if $b_{t+1i} > 0$.

In the case with period-by-period budget balance and no government debt, total capital in period t is $K_t = b_t$. Total labor is $L_t = \int_i v_{ti} l_{ti}$. Total product is $y_t = F(K_t, L_t)$ with CRS production function. Factor prices are given by $R_t = 1 + F_K$ and $w_t = F_L$ so that $F(K_t, L_t) = (R_t - 1)K_t + w_t L_t$.

The government objective is to choose $(\underline{R}_t, \underline{w}_t)_{t \geq 0}$ to maximize

$$SWF = \sum_{t \geq 0} \Delta^t \int_i \omega_{ti} V^{ti}(\underline{R}_t b_{ti} + \underline{w}_t v_{ti} l_{ti} + E_t - b_{t+1i}, \underline{R}_{t+1} b_{t+1i}, l_{ti}),$$

subject to the period-by-period budget constraint

$$E_t = (w_t - \underline{w}_t)L_t + (R_t - \underline{R}_t)b_t = b_t + F(b_t, L_t) - \underline{w}_t L_t - \underline{R}_t b_t.$$

Effectively, R_t and w_t have disappeared from the maximization problem. We can consider as above a small reform $(d\underline{R}_t = d\underline{R}, d\underline{w}_t)_{t \geq T}$ with $d\underline{w}_t$ set to meet the period-by-period budget constraint. The period by period budget constraint implies

$$L_t d\underline{w}_t + (w_t - \underline{w}_t)dL_t + b_t d\underline{R}_t + (R_t - \underline{R}_t)db_t = 0, \quad \text{so that}$$

$$b_t d\underline{R}_t \left(1 - e_{Bt}\frac{R_t - \underline{R}_t}{\underline{R}_t}\right) = -L_t d\underline{w}_t \left(1 - e_{Lt}\frac{w_t - \underline{w}_t}{\underline{w}_t}\right), \tag{A3}$$

where elasticities e_{Bt} and e_{Lt} are again defined with respect to $\underline{R}_t$ and $\underline{w}_t$ and hence are exactly equivalent to our earlier elasticities with respect to $1 - \tau_{Bt}$ and $1 - \tau_{Lt}$, i.e., they are pure supply elasticities keeping the pre-tax price of factors constant. Noting that $\frac{\tau_{Bt}}{1-\tau_{Bt}} = \frac{R_t - \underline{R}_t}{\underline{R}_t}$ and $\frac{\tau_{Lt}}{1-\tau_{Lt}} = \frac{w_t - \underline{w}_t}{\underline{w}_t}$, calculations follow exactly those from appendix A.1.1 and we obtain exactly the same formula (8).

In the case with government debt, the government dynamic budget constraint

$$a_{t+1} = R_t a_t + (R_t - \underline{R}_t)b_t + (w_t - \underline{w}_t)L_t - E_t$$

can be rewritten as

$$a_{t+1} = a_t + b_t + F(b_t + a_t, L_t) - \underline{R}_t b_t - \underline{w}_t L_t - E_t.$$

We can consider again the same small reform $(d\underline{R}_t = d\underline{R}, d\underline{w}_t)_{t \geq T}$ with $d\underline{w}_t$ set to meet the period-by-period budget constraint (A3) so that $da_t = 0$ for all t and the calculations are exactly as in the period-by-period budget balance case. This shows that formula (8) also applies with government debt.

A.1.3 Case with Economic Growth

We consider standard labor augmenting economic growth at rate $G > 1$ per generation, i.e., individual wage rates w_{ti} grow exogenously at rate G. Obtaining a steady-state where all variables grow at rate G per generation requires imposing standard homogeneity assumptions on individual utilities so that

$$V^{ti}(c, \underline{b}, l) = \frac{\left(U^{ti}(c, \underline{b})e^{-h_{ti}(l)}\right)^{1-\gamma}}{1-\gamma}$$

with $U^{ti}(c, \underline{b})$ homogeneous of degree one. In that case, the individual maximization problem can be decomposed into two steps.

First, the individual chooses b_{t+1i} taking resources $y_{ti} = Rb_{ti}(1 - \tau_{Bt}) + w_{ti}l_{ti}(1 - \tau_{Lt}) + E_t$ as given so that we can define the indirect utility:

$$v^{ti}(y_{ti}, R(1 - \tau_{Bt+1})) = \max_{b_{t+1i} \geq 0} U^{ti}(y_{ti} - b_{t+1i}, Rb_{t+1i}(1 - \tau_{Bt+1})).$$

Because U^{ti} is homogeneous of degree one, $v^{ti}(y, R(1-\tau_{Bt+1})) = y \cdot \phi^{ti}(R(1-\tau_{Bt+1}))$ is *linear* in y.

Second, the individual chooses labor supply to maximize

$$\log[\phi^{ti}(R(1-\tau_{Bt+1}))] + \log[Rb_{ti}(1-\tau_{Bt}) + w_{ti}(1-\tau_{Lt})l_{ti} + E_t] - h_{ti}(l_{ti}),$$

leading to the first order condition:

$$h'_{ti}(l_{ti}) = \frac{w_{ti}(1-\tau_{Lt})}{Rb_{ti}(1-\tau_{Bt}) + w_{ti}(1-\tau_{Lt})l_{ti} + E_t}$$

Hence, if tax rates converge and w_{ti}, b_{ti}, E_t, all grow at rate G per generation, labor supply l_{ti} will be stationary so that an ergodic equilibrium exists (under the standard assumptions).

This implies that utility V^{ti} grows at rate $G^{1-\gamma}$ per generation. As $V_c^{ti}/V^{ti} = (1-\gamma)/y_{ti}$ and y_{ti} grows at rate G, marginal utility V_c^{ti} grows at rate $G^{-\gamma}$ per generation.[43]

Steady-state maximization. Suppose that the government maximizes steady-state social welfare as in the main text. We obtain the same equation (5) as in the main text. However, the last term in b_{t+1i} has grown by a factor G relative to b_t so that when dividing (5) by $Rb_t d\tau_B$, we obtain:

$$0 = -\bar{b}^{\text{received}}(1+\hat{e}_B) + \frac{1-e_B\tau_B/(1-\tau_B)}{1-e_L\tau_L/(1-\tau_L)}\bar{y}_L - \frac{G\bar{b}^{\text{left}}}{R(1-\tau_B)},$$

which is the same equation as in the main text except that the term $\bar{b}^{\text{left}}$ is multiplied by a factor G. This will lead to the same optimum formula as (6) except that $\bar{b}^{\text{left}}$ is replaced by $G\bar{b}^{\text{left}}$, or equivalently R is replaced by R/G, i.e.,

$$\tau_B = \frac{1 - \left[1 - \frac{e_L\tau_L}{1-\tau_L}\right] \cdot \left[\frac{\bar{b}^{\text{received}}}{\bar{y}_L}(1+\hat{e}_B) + \frac{G}{R}\frac{\bar{b}^{\text{left}}}{\bar{y}_L}\right]}{1 + e_B - \left[1 - \frac{e_L\tau_L}{1-\tau_L}\right]\frac{\bar{b}^{\text{received}}}{\bar{y}_L}(1+\hat{e}_B)}. \tag{A4}$$

Social discounting maximization. We assume now that the government maximizes the discounted social welfare function:

$$SWF = \sum_{t\geq 0} \Delta^t \int_i \omega_{ti} V^{ti}(Rb_{ti}(1-\tau_{Bt}) + w_{ti}l_{ti}(1-\tau_{Lt}) + E_t - b_{t+1i}, Rb_{t+1i}(1-\tau_{Bt+1}), l_{ti}),$$

subject to period-by-period budget balance $E_t = \tau_{Bt}Rb_t + \tau_{Lt}y_{Lt}$. Consider again a reform $d\tau_B$ so that $d\tau_{Bt} = d\tau_B$ for all $t \geq T$ (and correspondingly $d\tau_{Lt}$ to maintain budget balance and keeping E_t constant). We assume that T is large enough that all variables have converged for $t \geq T$.

$$dSWF = \sum_{t\geq T} \Delta^t \int_i \omega_{ti} V_c^{ti} \cdot (Rdb_{ti}(1-\tau_{Bt}) - Rb_{ti}d\tau_B - d\tau_{Lt}y_{Lti}) + \sum_{t\geq T-1} \Delta^t \int_i \omega_{ti} V_{\underline{b}}^{ti} \cdot (-d\tau_B Rb_{t+1i}).$$

[43] This result remains true in the log-case with $\gamma = 1$.

We define elasticities e_{Bt} and e_{Lt} exactly as in equation (A1) in appendix A.1.1. We define $g_{ti} = \omega_{ti} V_c^{ti} / \int_j \omega_{tj} V_c^{tj}$ the normalized social marginal welfare weight on individual ti. Importantly, $\int_j \omega_{tj} V_c^{tj}$ now grows at rate $G^{-\gamma}$ per generation so that $G^{\gamma t} \int_j \omega_{tj} V_c^{tj}$ converges to a steady state.

Using the individual first order condition $V_c^{ti} = R(1-\tau_B) V_{\underline{b}}^{ti}$ when $b_{t+1i} > 0$, along with the budget balance equation (A1), and dividing by $G^{\gamma t} \int_j \omega_{tj} V_c^{tj}$ (constant in steady-state), allows to rewrite the first order condition $dSWF = 0$ as

$$0 = \sum_{t \geq T} \Delta^t G^{-\gamma t} \int_i g_{ti} \left[-R b_{ti}(1 + e_{Bti}) + \frac{1 - e_{Bt}\tau_B/(1-\tau_B)}{1 - e_{Lt}\tau_L/(1-\tau_L)} \frac{y_{Lti}}{y_{Lt}} R b_t \right] - \sum_{t \geq T-1} \Delta^t G^{-\gamma t} \int_i g_{ti} \frac{b_{t+1i}}{1-\tau_B},$$

As everything has converged for $t \geq T$, dividing by $R b_t G^{-t}$ (which is constant in steady-state) and denoting by $\bar{y}_L$, $\bar{b}^{\text{received}}$, $\bar{b}^{\text{left}}$, the population averages of $g_{ti} \cdot y_{Lti}/y_{Lt}$, $g_{ti} \cdot b_{ti}/b_t$, $g_{ti} \cdot b_{t+1i}/b_{t+1}$, and $\hat{e}_{Bt} = \int_i g_{ti} b_{ti} e_{Bti} / \int_i g_{ti} b_{ti}$, the first order condition is rewritten as:

$$0 = -\sum_{t \geq T} \Delta^t G^{t-\gamma t} \bar{b}^{\text{received}} (1 + \hat{e}_{Bt}) + \sum_{t \geq T} \Delta^t G^{t-\gamma t} \frac{1 - e_{Bt}\tau_B/(1-\tau_B)}{1 - e_{Lt}\tau_L/(1-\tau_L)} \bar{y}_L - \sum_{t \geq T-1} \Delta^t G^{t-\gamma t} \frac{G \bar{b}^{\text{left}}}{R(1-\tau_B)},$$

There are two differences with the case without growth. First, the G in the numerator of the last term appears because bequests left are from the next period and hence bigger by a factor G (exactly as in the steady-state maximization case presented above). Second, the discount factor Δ is replaced by $\Delta G^{1-\gamma}$ because of growth of all quantities (the G factor) and decrease in average marginal utility (the $G^{-\gamma}$ factor).

We define $e_B = (1 - \Delta G^{1-\gamma}) \sum_{t \geq T} (\Delta G^{1-\gamma})^{t-T} e_{Bt}$, $\hat{e}_B = (1 - \Delta G^{1-\gamma}) \sum_{t \geq T} (\Delta G^{1-\gamma})^{t-T} \hat{e}_{Bt}$ as the discounted average of the e_{Bt} and $\hat{e}_{Bt}$. We then define e_L so that:

$$\frac{1 - e_B\tau_B/(1-\tau_B)}{1 - e_L\tau_L/(1-\tau_L)} = (1 - \Delta G^{1-\gamma}) \sum_{t \geq T} (\Delta G^{1-\gamma})^{t-T} \frac{1 - e_{Bt}\tau_B/(1-\tau_B)}{1 - e_{Lt}\tau_L/(1-\tau_L)}.$$

Naturally, in the case e_{Lt} constant in t, then we have $e_{Lt} \equiv e_L$. Using those definitions we can rewrite the first order condition as:

$$0 = -\bar{b}^{\text{received}}(1 + \hat{e}_B) + \frac{1 - e_B\tau_B/(1-\tau_B)}{1 - e_L\tau_L/(1-\tau_L)} \bar{y}_L - \frac{G \bar{b}^{\text{left}}}{R \Delta G^{1-\gamma}(1-\tau_B)},$$

where the $\Delta G^{1-\gamma}$ expression in the denominator of the third term appears because the sum for the third term starts at $T-1$ instead of T. Re-arranging this expression leads immediately to formula (8) with Δ being replaced by $\Delta G^{-\gamma}$, i.e.,

$$\tau_B = \frac{1 - \left[1 - \frac{e_L\tau_L}{1-\tau_L}\right] \cdot \left[\frac{\bar{b}^{\text{received}}}{\bar{y}_L}(1 + \hat{e}_B) + \frac{1}{R\Delta G^{-\gamma}} \frac{\bar{b}^{\text{left}}}{\bar{y}_L}\right]}{1 + e_B - \left[1 - \frac{e_L\tau_L}{1-\tau_L}\right] \frac{\bar{b}^{\text{received}}}{\bar{y}_L}(1 + \hat{e}_B)}, \qquad \text{(A5)}$$

When the modified Golden rule holds, we have $R\Delta G^{-\gamma} = 1$ so that formula (9) applies unchanged (all the reasoning with endogenous capital stock applies virtually unchanged). The proof of the Modified Golden Rule with growth can be done exactly as in the case with no growth by considering one small reform $d\underline{w}$ at period T and the same reform (multiplied by $-R$) at period $T+1$. By linearity of small changes, the sum of the two reforms is budget neutral. Hence, it has to be welfare neutral as well. The social welfare effect of the period $T+1$ reform is $-R\Delta G^{-\gamma}$ times the welfare effect of the period T reform because (a) it is $-R$ times bigger, (b) it happen one generation later so is discounted by Δ, (c) it affects generations which have marginal utility $G^{-\gamma}$ times as large.

A.1.4 Anticipated and Long-Run Elasticities in the Dynastic Model

In this section, we want to provide detailed intuitions for why the anticipatory elasticity $e_B^{\text{anticip.}}$, the post-reform elasticity e_B^{post} and the long-run steady-state elasticity e_B are all finite in the ergodic model with stochastic wages (the Aiyagari model) and why they become infinite when stochastic shocks vanish (the Chamley-Judd model). Last, we want to show that, in the Chamley-Judd model with endogenous discount rate, the long-run elasticity e_B may be finite but the anticipatory elasticity is still infinite. We only provide intuitions rather than fully rigorous detailed proofs because the formal proof can be immediately obtained by combining the Chamley-Judd and Aiyagari results with our optimal tax formulas.[44]

Let us consider as in the text, a reform $d\tau_{Bt} = d\tau_B$ for $t \geq T$ in the distant future. This reform generates aggregate changes in bequests db_t so that we can define period t elasticity as $e_{Bt} = -\frac{1-\tau_{Bt}}{b_t}\frac{db_t}{d\tau_B}$. Recall that elasticities are defined as

$$e_B^{\text{post}} = (1-\delta)\sum_{t\geq T}\delta^{t-T}e_{Bt}, \quad e_B^{\text{anticip.}} = (1-\delta)\sum_{t<T}\delta^{t-T}e_{Bt}, \quad \text{and} \quad e_B = \lim_{t\to\infty} e_{Bt} \tag{A6}$$

Non stochastic wages (Chamley-Judd). Let us consider first the standard case with uniform and constant discount rate δ as in the main text. Let us further assume that $\delta R = 1$ and that $\tau_{Bt} \equiv 0$ so that we start from an initial situation with a well defined steady-state.

In the Chamley-Judd model, future wages y_{Lti} are fully known as of period zero. In that case, the natural assumption is that there are no credit constraints and hence the individual first order condition $u'(c_{ti}) = \delta R(1-\tau_{Bt+1})u'(c_{t+1i})$ always holds.[45]

In that case, with $\delta R = 1$ and $\tau_{Bt} \equiv 0$, the individual fully smoothes consumption $c_{ti} = c_{0i}$ for all t with $c_{0i} = [1-1/R](b_{0i} + \sum_{t\geq 0} y_{Lti}/R^t)$ to satisfy the inter-temporal budget.

The future tax reform leads to a decreasing consumption path after the reform and a flat shift of the pre-reform consumption path proportional to $R^{-T}d\tau_B$ (as it affects the PDV of resources by a factor proportional to $R^{-T}d\tau_B$). The aggregated budget constraint implies that $b_{t+1} = Rb_t + y_{Lt} - c_t$ for $t < T$ so that $b_t = R^t b_0 - c_0[1 + R + .. + R^{t-1}] + y_{Lt-1} + .. + R^{t-1}y_{L0}$

[44]Namely, the fact that Chamley-Judd obtain a zero optimal long-run rate implies that the elasticity e_B^T is infinite. The fact that Aiyagari obtains a positive optimal long-run rate implies that the elasticity e_B^T is finite.

[45]This will be true for large t without any assumption if y_{Lti} converges to a constant y_{Li} for large t, the natural assumption for steady-state reasoning.

and hence $db_t = -dc_0(R^t - 1)/(R-1) \sim R^{t-T} d\tau_B$. This implies that $e_{Bt} \sim R^{t-T}$. Therefore, $e_B^{\text{anticip.}} = (1-\delta)\sum_{t<T} \delta^{t-T} e_{Bt} \sim (1-\delta)\sum_{t<T}(\delta R)^{t-T} = (1-\delta)T$ is infinite for large T.

As is well known, the long-run elasticity e_B is also infinite as any long-run tax starting from a $\delta R = 1$ steady-state leads to an exponentially decreasing path of consumption and hence as much individual debt as possible.

Let us know consider the case with endogenous discount factor $\delta^i(c)$, decreasing in c. In that case, there is a steady-state such that $\delta^i(c_i)R = 1$ for all i. Intuitively, if $\delta^i R > 1$, individual i accumulates more wealth, eventually allowing him to consume more so that $\delta^i R$ is driven down to one (and conversely). In that case in steady-state (when all variables have converged), $c_i = (R-1)b_i + y_{Li}$ and b_i is an implicit function of R through the equation $\delta^i((R-1)b_i + y_{Li})R = 1$. Hence, the individual supply b_i is a smooth function of R. Hence aggregate long run bequests b are also a smooth function of R and the long-run elasticity e_B is therefore finite.

In that case however, it is still the case that a future reform is going to shift the entire (pre-reform) consumption path so that $dc_{0i} \sim R^{-T} d\tau_B$, which implies $db_t \sim R^{t-T} d\tau_B$, $e_{Bt} \sim R^{t-T}$ Therefore, $e_B^{\text{anticip.}} \sim T$ is infinite for large T implying that the optimal long-run tax rate τ_B is zero in spite of a finite long-run elasticity e_B, an important point made by Judd (1985), Theorem 5.

Stochastic wages (Aiyagari). In the stochastic model, individual ti solves the problem:

$$\max_{b_{t+1i}} u(Rb_{ti}(1-\tau_{Bt}) + \tau_{Bt}Rb_t + y_{Lti} - b_{t+1i}) + \sum_{s>t} \delta^{s-t} E_t u(Rb_{si}(1-\tau_{Bs}) + \tau_{Bs}Rb_s + y_{Lsi} - b_{s+1i}).$$

In any path where $b_{si} = 0$ for some s such that $t < s \leq T$, any current marginal bequest change db_{t+1i} has zero impact on post s generations and hence the future tax rate change $(d\tau_{Bt})_{t \geq T}$ is irrelevant for the current decision b_{t+1i}. Concretely, in the scenario where my child is going to fully consume my bequest and leave nothing to my grandchildren, a marginal increase in bequest taxes for my grand-children does not affect my own bequest decision. Hence, the behavioral response db_{t+1i} to the future tax increase is discounted relative to the Chamley-Judd model with no uncertainty by a factor $\kappa^{ti}(T-(t+1))$ which is the probability that all my descendants from $s = t+1$ to $s = T$ all leave positive bequests $b_{si} > 0$.

$\kappa^{ti}(T-(t+1))$ is of course higher if I am very wealthy as my bequest is less likely to get fully consumed quickly by my descendants. By ergodicity, as my initial wealth is irrelevant in the distant future, it must be the case that for $T-s$ large, $\kappa^{ti}(T-(s+1))/\kappa^{ti}(T-s)$ converges to a constant $\kappa < 1$ that depends on the structure of shocks, the tax system, etc. but is uniform across individuals. This constant κ is equal to the fraction of individuals with positive bequests in the ergodic cross-section. Hence, κ is necessarily less than one as long as the fraction of individuals with zero bequests is strictly positive in steady-state. Naturally, when uncertainty in future labor shocks vanishes, κ converges to one.

Hence, at the aggregate level, the response db_t to the future tax increase starting at date T is reduced by an exponential factor proportional to κ^{T-t}.

As we have seen, in the Chamley-Judd model with no uncertainty and $\delta R = 1$, we have $db_t \sim \delta^{T-t} d\tau_B$. Hence, with stochastic shocks, $db_t \sim \delta^{T-t}\kappa^{T-t} d\tau_B$ so that $e_{Bt} \sim \delta^{T-t}\kappa^{T-t}$. This implies that $e_B^{\text{anticip.}} = (1-\delta)\sum_{t<T} \delta^{t-T} e_{Bt} \sim (1-\delta)\sum_{t<T} \kappa^{T-t} = (1-\delta)/(1-\kappa)$ is finite.

In the ergodic long-run, with stochastic shocks, aggregate bequests b_t converge so naturally b_t will be a smooth function of τ_B.[46] In that case, the long-run elasticity e_B is finite. This also implies that the post-elasticity $e_B^{\text{post}} = (1-\delta)\sum_{t\geq T} \delta^{t-T} e_{Bt}$ is finite, which establishes that $e_B^T = e_B^{\text{anticip.}} + e_B^{\text{post}}$ is finite and delivers a non-zero optimal τ_B as in Aiyagari (2005).

A.1.5 Formula (14) in Dynastic Model with Elastic Labor Supply

We consider the small open economy with exogenous R, period-by-period budget balance, and the utilitarian case. The government chooses $(\tau_{Bt}, \tau_{Lt})_{t\geq 0}$ to maximize

$$EV_0 = \sum_{t\geq 0} \delta^t E u^{ti}(Rb_{ti}(1-\tau_{Bt}) + (1-\tau_{Lt})w_{ti}l_{ti} + E_t - b_{t+1i}, l_{ti}).$$

subject to period-by-period budget balance $E_t = \tau_{Bt}Rb_t + \tau_{Lt}y_{Lt}$ with E_t given (and converging to E). Note that we allow for heterogeneity in felicity functions $u^{ti}(c,l)$ as this does not affect the analysis.

Consider again a reform $d\tau_B$ so that $d\tau_{Bt} = d\tau_B$ for all $t \geq T$ (and correspondingly $d\tau_{Lt}$ to maintain budget balance and keeping E_t constant). We assume that T is large enough that all variables have converged for $t \geq T$. Using the envelope condition as l_{ti} and b_{ti} maximize individual utility, we have:

$$0 = dEV_0 = -\sum_{t\geq T} \delta^t E[u_c^{ti}(c_{ti}, l_{ti})Rb_{ti}]d\tau_{Bt} - \sum_{t\geq 0} \delta^t E[u_c^{ti}(c_{ti}, l_{ti})d\tau_{Lt}y_{Lti}]$$

To rewrite this equation in terms of elasticities of b_t and y_{Lt} with respect to $1-\tau_B$ and $1-\tau_L$, we can define e_{Bt} as the elastic response of b_t to the tax reform $d\tau = (d\tau_{Bt}, d\tau_{Lt})_{t\geq 0}$, so that $\frac{db_t}{b_t} = -e_{Bt}\frac{d\tau_B}{1-\tau_B}$ where db_t is the aggregate bequest response to the full reform $d\tau$. Note that the response of b_t may start before period T due to anticipatory effects described in the text. Such anticipatory effects start before T but are vanishingly small as distance to the reform increases. Therefore, we can assume that anticipatory effects take place only after all variables have converged (as long as T is chosen large enough).

The response builds over generations and eventually converges to the long-run steady-state elasticity e_B. We similarly define the elasticity e_{Lt} so that $\frac{dy_{Lt}}{y_{Lt}} = -e_{Lt}\frac{d\tau_{Lt}}{1-\tau_L}$ where dy_{Lt} is the labor supply response to the full reform $d\tau$. Period-by-period budget balance requires:

$$Rb_t d\tau_B \left(1 - e_{Bt}\frac{\tau_B}{1-\tau_B}\right) = -d\tau_{Lt}y_{Lt}\left(1 - e_{Lt}\frac{\tau_L}{1-\tau_L}\right) \quad \text{for} \quad t \geq T,$$

[46] For $\tau_B = 1$, there is no incentive to leave bequests and $b_t = 0$. Conversely, for sufficiently large subsidies, if $\delta R(1-\tau_B) > 1$ then bequests b_t would explode. In between, b_t is a smooth function of $R(1-\tau_B)$.

$$-Rb_t d\tau_B e_{Bt}\frac{\tau_B}{1-\tau_B} = -d\tau_{Lt} y_{Lt}\left(1-e_{Lt}\frac{\tau_L}{1-\tau_L}\right) \quad \text{for} \quad t<T.$$

The equation for $t<T$ does not have the term $Rb_t d\tau_B$ on the left-hand-side because the $d\tau_B$ reform starts at T. However, through anticipatory responses, b_t responds before T, requiring an adjustment $d\tau_{Lt}$ to balance the budget (and which triggers a labor supply response). Using those equations, (and dividing by $Rb_t d\tau_B$ as b_t is constant in the long-term), we can rewrite $dEV_0 = 0$ as:

$$0 = -\sum_{t\geq T}\delta^t E\left[u_c^{ti}\frac{b_{ti}}{b_t}\right] + \sum_{t\geq T}\delta^t E\left[u_c^{ti}\frac{y_{Lti}}{y_{Lt}}\right]\frac{1-\frac{e_{Bt}\tau_B}{1-\tau_B}}{1-\frac{e_{Lt}\tau_L}{1-\tau_L}} - \sum_{t<T}\delta^t E\left[u_c^{ti}\frac{y_{Lti}}{y_{Lt}}\right]\frac{\frac{e_{Bt}\tau_B}{1-\tau_B}}{1-\frac{e_{Lt}\tau_L}{1-\tau_L}}.$$

As everything has converged, denoting by $\bar{y}_L$, $\bar{b}^{\text{received}}$ the population averages of $u_c^{ti}\cdot y_{Lti}/[y_{Lt}Eu_c^{ti}]$, $u_c^{ti}\cdot b_{ti}/[b_t Eu_c^{ti}]$, we get:

$$0 = -\bar{b}^{\text{received}}\sum_{t\geq T}\delta^t + \bar{y}_L\sum_{t\geq T}\delta^t\frac{1-\frac{e_{Bt}\tau_B}{1-\tau_B}}{1-\frac{e_{Lt}\tau_L}{1-\tau_L}} - \bar{y}_L\sum_{t<T}\delta^t\frac{\frac{e_{Bt}\tau_B}{1-\tau_B}}{1-\frac{e_{Lt}\tau_L}{1-\tau_L}}.$$

Defining the bequest elasticities as in the main text

$$e_B^T = e_B^{\text{post}} + e_B^{\text{anticip.}} \quad \text{with } e_B^{\text{post}} = (1-\delta)\sum_{t\geq T}\delta^{t-T}e_{Bt} \quad \text{and } e_B^{\text{anticip.}} = (1-\delta)\sum_{t<T}\delta^{t-T}e_{Bt},$$

and defining e_L so that:

$$\frac{1-e_B^T\tau_B/(1-\tau_B)}{1-e_L\tau_L/(1-\tau_L)} = (1-\delta)\sum_{t\geq T}\delta^{t-T}\frac{1-e_{Bt}\tau_B/(1-\tau_B)}{1-e_{Lt}\tau_L/(1-\tau_L)} - (1-\delta)\sum_{t<T}\delta^{t-T}\frac{e_{Bt}\tau_B/(1-\tau_B)}{1-e_{Lt}\tau_L/(1-\tau_L)}.$$

Again, in the case e_{Lt} constant in t, then we have $e_{Lt}\equiv e_L$. Such would be the case with iso-elastic quasi-linear utility functions of the form $V^{ti}(c,\underline{b},l) = U^{ti}(c-l^{1+1/e_L},\underline{b})$ where labor supply depends solely on the net-of-tax wage rate and has no income effects. Using those definitions we can rewrite the first order condition as:

$$0 = -\bar{b}^{\text{received}} + \bar{y}_L\frac{1-e_B^T\tau_B/(1-\tau_B)}{1-e_L\tau_L/(1-\tau_L)}.$$

This can be easily re-arranged in the first formula in (14).

To obtain the second formula, we aggregate the individual first order condition

$$u_c^{ti}(c_{ti},l_{ti})b_{t+1i} = \delta R(1-\tau_{Bt+1})b_{t+1i}E_t[u_c^{t+1i}(c_{t+1i},l_{t+1i})]$$

so that $\bar{b}^{\text{left}} = \delta R(1-\tau_B)\bar{b}^{\text{received}}$ in the long-run steady-state with $\bar{b}^{\text{left}}$ the population average of $u_c^{ti}\cdot b_{t+1i}/[b_t Eu_c^{ti}]$. Q.E.D.

A.1.6 Modified Golden Rule in the Dynastic Model

We can extend the dynastic model to the case with endogenous factor prices (closed economy) exactly as our model of Section 3.1. Again, this extension requires to be able to tax both labor income and capital at separate and time varying rates so that the government controls after-tax factor prices $\underline{R}_t$ and $\underline{w}_t$. This implies that the optimal τ_B formula carries over to the close economy case unchanged. The formula applies both in the period by period budget balance case and when the government can use debt.

When the government can use debt optimally, the modified Golden rule $\delta R = 1$ holds also in the dynastic model. This can be established exactly in the same way as in our model of Section 3.1. We consider a small reform $d\underline{w}$ at period T and the same reform (multiplied by $-R$) at period $T+1$. By linearity of small changes, the sum of the two reforms is budget neutral. Hence, it has to be welfare neutral as well. The social welfare effect of the period $T+1$ reform is $-R\delta$ times the welfare effect of the period T reform because (a) it is $-R$ times bigger, (b) it happens one generation later so is discounted by δ. This implies that $\delta R = 1$. Note that Aiyagari (2005) obtains the same result but uses the government provided public good to establish it. Our proof shows that it is not necessary to have a public good. Any type of reform at periods T vs. $T+1$ can prove the result. This shows that the Modified Golden Rule is a very robust result of dynamic efficiency and optimal capital accumulation.

A.2 Optimal Taxation in the Overlapping Generation Model

A.2.1 The OLG Model and Optimal Tax Formula

The OLG model can be seen as a special case of our general model from Section 2 as follows. Each dynasty is a succession of (non-overlapping) two-period long lives with no altruistic linkage across lives within a dynasty. For example $i0$ is young person, $i1$ is the same person old, $i2$ is new young person altruistically unrelated to $i1$, etc. Cohort t denotes the individuals born in period t (young in period t and old in period $t+1$). Hence, the young receive no bequests, work, consume, and leave bequests to their future old self. The old do not work and simply consume their bequests. The preferences of the young take the form $V^{ti}(c_{ti}, Rb_{t+1i}(1-\tau_{Bt+1}), l_{ti})$. The preferences of the old are the same except that they have no choice left whatsoever and simply consume $Rb_{t+1i}(1-\tau_{Bt+1})$. To fit our model in our previous setting, we can for example assume that the utility of the young is separable in c and $\underline{b}$ so that $V^{ti}(c, \underline{b}, l) = u_{ti}(c) + v_{ti}(\underline{b}) - h_{ti}(l)$ and assume that the utility of the old is $v_{ti}(c)$.[47] In general equilibrium, this is effectively an OLG model with overlapping generations (assuming that young and old are in equal proportion in any cross-section). In that model, we can restrict the demogrant E_t to the young only.[48] The

[47] The linkage of the young and the old is necessarily altruistic in this basic model as the utility of the young can be always be rewritten as $V^{ti}(c_{ti}, u_{ti}, l_{ti})$.

[48] Demogrants to the old can be considered as well and eliminated without loss of generality by altering the young preferences so that they take into account the demogrant they will receive when old, i.e., their utility is

young solve the following maximization problem

$$\max_{c_{ti}, l_{ti}, b_{t+1i} \geq 0} V^{ti}(c_{ti}, Rb_{t+1i}(1-\tau_{Bt+1}), l_{ti}) \text{ s.t. } c_{ti} + b_{t+1i} = w_{ti} l_{ti}(1-\tau_{Lt}) + E_t. \tag{A7}$$

The natural social welfare criterion in this context is to put weight solely on the welfare of the young (as the utility of the old is taken into account through the utility of the young). In that case, we can immediately apply our optimal previous formula (9) in the dynamically efficient case to obtain the following formula:

$$\textbf{Optimal tax on savings in OLG model} \quad \tau_B = \frac{1 - \left[1 - \frac{e_L \tau_L}{1-\tau_L}\right] \cdot \frac{\bar{b}^{\text{left}}}{\bar{y}_L}}{1 + e_B}. \tag{A8}$$

This formula naturally assumes that $\bar{b}^{\text{received}} = 0$ as bequest receivers (the old) have no separate weighting in the social welfare function. The formula also assumes that we use social discounting Δ at the Modified Golden Rule so that $R\Delta = 1$. As we show in appendix A.3, the formula can actually be obtained without needing to refer to social discounting or the Modified Golden Rule by considering steady-state welfare maximization subject to the "generational" budget balance $E_t = \tau_{Bt} b_{t+1} + \tau_L y_{Lt}$ instead of the cross-sectional budget balance $E_t = \tau_{Bt} R b_t + \tau_L y_{Lt}$.

A.2.2 The Zero-Tax Result

We can therefore consider steady-state social welfare maximization subject to the generational budget constraint $E_t = \tau_{Bt} b_{t+1} + \tau_L y_{Lt}$ and hence drop any t subscripts in what follows. As in the text, we now specialize utility functions to such that $V^i(c, \underline{b}, l) = U^i(u(c, \underline{b}), l)$ with the sub-utility of consumption $u(c, \underline{b})$ homogeneous of degree 1 and homogeneous across individuals in the population.

Let us prove that any budget neutral tax system (τ_B, τ_L, E) can be replaced by an alternative tax system $(\tau_B = 0, \tau_L', E')$ that leaves everybody's utility unchanged and raises at least as much revenue. We adapt the Kaplow (2006) and Laroque (2005) recent and elegant proof of the Atkinson-Stiglitz theorem to the linear earnings tax case.

Let us denote by $p = \frac{1}{R}$ and $q = \frac{1}{R(1-\tau_B)}$ the pre-tax and post-tax prices of period 2 consumption $\underline{b}$. Let

$$v(y, q) = \max_{c, \underline{b} \geq 0} u(c, \underline{b}) \quad \text{s.t.} \quad c + q\underline{b} \leq y,$$

be the indirect utility of consumption. Because $u(c, \underline{b})$ is homogeneous of degree one, $v(y, q)$ is linear in y to that $v(y, q) = y \cdot \phi(q)$ as we showed in the case with economic growth in appendix A.1.3.

Starting from the initial tax system (τ_B, τ_L, E), let us consider the alternative tax system $(\tau_B = 0, \tau_L', E')$ such that $\phi(p)(1-\tau_L') = \phi(q)(1-\tau_L)$ and $\phi(p)E' = \phi(q)E$. This alternative system is precisely designed so that $v(y_{Li}(1-\tau_L') + E', p) = v(y_{Li}(1-\tau_L) + E, q)$ for all y_{Li}. Hence,

---

$V^{ti}(c_{ti}, Rb_{t+1i}(1-\tau_{Bt+1}) + E^{\text{old}}_{t+1}, l_{ti})$.

it leaves all individual utilities and labor supply choices unchanged. Finding this alternative linear tax system is feasible because $v(y,q) = y \cdot \phi(q)$ is linear in y.

Let us now show that the alternative tax system raises as much revenue as the initial tax system. Suppose individual i chooses $(l_i, c_i, \underline{b}_i)$ under the initial tax system so that $c_i + q\underline{b}_i = w_i l_i(1-\tau_L) + E$. Attaining utility $v(y_{Li}(1-\tau'_L) + E', p)$ with $\tau'_B = 0$ costs $y_{Li}(1-\tau'_L) + E'$. As $v(y_{Li}(1-\tau_L) + E, q) = v(y_{Li}(1-\tau'_L) + E', p)$, this utility is also attained under the initial choice $(l_i, \underline{b}_i)$ which must henceforth cost at least $y_{Li}(1-\tau'_L)+E'$ under price p so that $y_{Li}(1-\tau'_L)+E' \leq c_i + p\underline{b}_i = y_{Li}(1-\tau_L) + E - \underline{b}_i(q-p)$. This implies that $\tau_L y_{Li} + (q-p)\underline{b}_i \leq \tau'_L y_{Li}$, i.e., $\tau_L y_{Li} + \tau_B b_i \leq \tau'_L y_{Li}$, so that the new tax system raises at least as much revenue individual by individual.

The key lesson from this analysis is that zero capital tax is desirable in the OLG model because any capital tax can be replaced by a more efficient labor tax but this is possible only because labor earnings is the single dimension of inequality in the OLG model.

A.3 Pedagogical Rawlsian Meritocratic Optimal Formulas

In the case of the Rawlsian Meritocratic optimum where social welfare is concentrated among zero-receivers, it is possible to obtain the long-run optimum tax formula (9) that maximizes discounted social welfare with dynamic efficiency as the solution of the much simpler following static problem. The government maximizes steady-state welfare subject to the alternative "generational" budget balance $\tau_{Bt} b_{t+1} + \tau_{Lt} y_{Lt} = E_t$. The idea is that generation t funds its demogrant E_t with taxes on its labor earnings y_{Lt} and taxes on the bequests it leaves. Bequest taxes are collected at the end of the period.[49] This derivation is useful because it delivers the Meritocratic Rawlsian version of (9) without having to worry about discounting and dynamic efficiency issues, and hence can be particularly useful for teaching purposes.

Formally, assuming everything has converged to the steady-state (so that t subscripts can be dropped), the government maximizes

$$SWF = \max_{\tau_L, \tau_B} \int_i \omega_i V^i(w_i l_i(1-\tau_L) + E - b_i, R b_i(1-\tau_B), l_i) \quad s.t. \quad \tau_B b + \tau_L y_L = E. \tag{A9}$$

Note that bequests received are not included in life-time ressources because ω_i is zero for bequest receivers. We denote by $g_i = \omega_i V^i_c / \int_j \omega_j V^j_c$ the normalized social marginal welfare weight on individual i. g_i measures the social value of increasing consumption of individual i by \$1 (relative to increasing everybody's consumption by \$1).

Consider a small reform $d\tau_B > 0$, budget balance with $dE = 0$ requires that $d\tau_L$ is such that:

$$b d\tau_B \left(1 - e_B \frac{\tau_B}{1-\tau_B}\right) = -d\tau_L y_L \left(1 - e_L \frac{\tau_L}{1-\tau_L}\right), \tag{A10}$$

[49] This is equivalent to collecting them on capitalized bequests Rb_{t+1} at the end of next period and discounting those taxes at rate $1/R$ as they accrue one period later.

where we have used the standard elasticity definitions (3).

Using the fact that b_i and l_i are chosen to maximize individual utility and applying the envelope theorem, the effect of the reform $d\tau_B, d\tau_L$ on steady-state social welfare is:

$$dSWF = \int_i \omega_i V_c^i \cdot (-d\tau_L y_{Lti}) + \omega_i V_{\underline{b}}^i \cdot (-d\tau_B R b_i).$$

At the optimum, $dSWF = 0$. Using the individual first order condition $V_c^i = R(1-\tau_B)V_{\underline{b}}^i$ when $b_i > 0$, expression (A10) for $d\tau_L$, and the definition of g_i, we have:

$$0 = \int_i g_i \cdot \left(\frac{1 - e_B\tau_B/(1-\tau_B)}{1 - e_L\tau_L/(1-\tau_L)} \frac{y_{Li}}{y_L} b d\tau_B - d\tau_B \frac{b_i}{1-\tau_B} \right),$$

The first term captures the positive effect of reduced labor income tax and the second term captures the negative effect on bequest leavers.

Let $\bar{y}_L$ and $\bar{b}^{\text{left}}$ be the population averages of $g_i \cdot y_{Li}/y_L$ and $g_i \cdot b_i/b$. The first order condition can therefore be rewritten as:

$$0 = \frac{1 - e_B\tau_B/(1-\tau_B)}{1 - e_L\tau_L/(1-\tau_L)} \bar{y}_L - \frac{\bar{b}^{\text{left}}}{1-\tau_B}, \quad \text{hence}$$

Meritocratic Rawlsian Steady-State and Dynamically Efficient Optimum. *The optimal tax rate τ_B that maximizes long-run welfare of zero-bequest receivers with period-by-period "generational" budget balance $\tau_{Bt} b_{t+1} + \tau_{Lt} y_{Lt} = E_t$ is given by:*

$$\tau_B = \frac{1 - \left[1 - \frac{e_L\tau_L}{1-\tau_L}\right] \cdot \frac{\bar{b}^{\text{left}}}{\bar{y}_L}}{1 + e_B}. \tag{A11}$$

Two points are worth noting about (A11).

First, this formula is consistent with the dynamically efficient formula because it considers the "generational" budget constraint $\tau_{Bt} b_{t+1} + \tau_{Lt} y_{Lt} = E_t$ instead of the cross-sectional budget constraint $\tau_{Bt} R b_t + \tau_{Lt} y_{Lt} = E_t$. This works for zero-receivers because the welfare trade-off involves solely current labor taxes vs. taxes paid on bequests left for the *same* generation t. If the social welfare function puts weight on bequests receivers, this "generational" budget fails to be consistent with the dynamic efficient case because of the welfare term involving bequests received.[50] In contrast the cross-sectional budget works for the term involving bequests received but fails for bequests left. Hence in the general case involving both bequests receivers and bequests leavers in the social welfare calculation, two generations are involved and there is no steady-state budget short-cut that can be consistent with the dynamically efficient case. In that case, we need to go back to the analysis presented in the main text.

[50] This term will be blown up by a factor R when using the generational budget. When discounting welfare with discount rate Δ, the blown up factor becomes $R\Delta$, which disappears when the Modified Golden Rule $R\Delta = 1$ holds.

Second, this formula can also be applied to the OLG model (where all welfare weights are naturally concentrated on zero-receivers as discussed in appendix A.2). Hence, the OLG analysis can also be done simply by considering the "generational" budget and by-passing entirely dynamic efficiency issues. This is well-known in the OLG Ramsey tax literature (see e.g., Atkinson and Sandmo 1980 and King 1980).

A.4 Link with Farhi and Werning (2010)

A.4.1 Fahri and Werning (2010) two-period model

Farhi and Werning (2010), Sections II-V consider a two-period model where parents do not have any initial wealth, work, and leave bequests to their children (who have no labor income and rely solely on such bequests). Inequality is due solely to heterogeneity in parent's earning abilities. Hence, this model is isomorphic to the OLG model where the young and the old are relabeled as two distinct individuals–namely parent and child. Farhi and Werning (2010) consider fully general nonlinear taxation. With standard "OLG" social welfare functions putting all the weight on the parent (and no weight on the children, i.e. the old of the conventional OLG interpretation), the Atkinson and Stiglitz theorem applies (under the standard weak separability assumption on preferences) so that there should be no marginal tax on bequests. As we discussed in section 3.2 and appendix A.2, in our linear tax model, if we impose the stronger homogeneity of degree one assumption on preferences for consumption on top of the standard Atkinson-Stiglitz weak separability assumption, we also obtain the same zero-bequest tax result.

Farhi and Werning (2010) then generalize the analysis to put additional weight on the utility of children. In that case, bequests have additional social value, and hence the zero-tax standard result is transformed into a negative marginal bequest tax result.

This result has a simple counter-part in our linear tax model. If we put weight on bequest receivers, (i.e., the non working old in the OLG interpretation which are the non-working children in the Farhi-Werning model) in our OLG set-up, formula (A8) is changed to

Optimal tax on bequests in the two-period Farhi-Werning linear tax model

$$\tau_B = \frac{1 - \left[1 - \frac{e_L \tau_L}{1-\tau_L}\right] \cdot \left[\frac{(1+\hat{e}_B)\bar{b}^{\text{received}}}{\bar{y}_L} + \frac{\bar{b}^{\text{left}}}{\bar{y}_L}\right]}{1 + e_B - \left[1 - \frac{e_L \tau_L}{1-\tau_L}\right] \frac{(1+\hat{e}_B)\bar{b}^{\text{received}}}{\bar{y}_L}} = \frac{-\left[1 - \frac{e_L \tau_L}{1-\tau_L}\right] \cdot \frac{\bar{b}^{\text{received}}}{\bar{y}_L}(1+\hat{e}_B)}{1 + e_B - \left[1 - \frac{e_L \tau_L}{1-\tau_L}\right] \frac{\bar{b}^{\text{received}}}{\bar{y}_L}(1+\hat{e}_B)} < 0, \quad \text{(A12)}$$

where the second equality holds whenever the standard OLG tax rate from formula (A8) is zero (i.e., under the homogeneity and separability assumptions discussed above). Hence our linear tax model is fully consistent with the Farhi-Werning two-period nonlinear tax analysis.

As our paper analysis has made clear, this result only provides an incomplete characterization of the bequest tax problem because it fails to capture the fact that lifetime resources inequality is bi-dimensional. In the Farhi-Werning two-period model, inequality is one-dimensional and hence there is no reason to tax bequests to start with. Adding a concern for bequests receivers then leads to negative bequests taxes. Farhi and Werning (2010) analysis does capture the important fact that bequest taxation hurts both bequest leavers and bequest receivers so that indeed, consistently with our analysis, for some social welfare functions, this double welfare cost of bequest taxation may lead to optimal negative bequest taxes (a point informally made by Kaplow, 2001).[51]

A.4.2 Fahri and Werning (2010) New Dynamic Public Finance model

Farhi and Werning then extend their two-period model to cast it into a full-fledged New Dynamic Public Finance (NDPF) infinite horizon model where individuals are both bequests receivers and bequest leavers. In the standard case where welfare is measured solely from period 0, the standard NDPF inverse Euler equation applies, so that there is a positive inter-temporal tax wedge, i.e., an implicit tax rate on bequests. As shown by Kocherlakota (2005) however, the optimal mechanism can also be decentralized with stochastic linear bequest taxes (whose rate depends on the full history of shocks and also receivers wages yet unknown at time of bequest), where the expected inheritance tax rate is zero for all individuals. Again, this nominally provides a zero-tax benchmark although the economic reasons for the zero-tax in this dynamic model are drastically different from simple zero-tax intuition of the Atkinson-Stiglitz two-period type model.

First, the zero-tax result of Kocherlakota (2005) depends on the decentralization mechanism chosen. Other forms of decentralization that have been proposed in earlier NDPF papers with more specific wage shock structures (such as iid shocks) can lead to decentralization with positive capital taxes (as in the initial interpretation of the inverse Euler equation). A striking feature of the Kocherlakota (2005) implementation presented in Farhi and Werning (2010), p. 664, is that everybody leaves exactly the same bequest at the optimum so that the conventional notion of wealth inequality disappears, making it difficult to map this theoretical result into actual policy.[52]

Second and more important, as pointed out in Farhi and Werning (2007), the Kocherlakota optimum has no steady-state as inequality becomes more and more skewed (the so-called immiseration result) where everybody's consumption goes to zero almost surely. In our view, this is a severe limitation of the framework, which makes it difficult to use for practical policy recommendations.[53] In contrast, in our linear tax model, a steady-state exists and the long-run optimal linear bequest tax τ_B is positive (consistent with the earlier analysis of Aiyagari, 2005).

When some weight is put on future generations (either through social discounting that is

[51] Farhi and Werning (2010) consider the fully nonlinear case and obtain valuable results on the progressivity of the optimal bequest tax subsidy that cannot be captured in our linear framework.

[52] Finding a NDPF implementation that retains a more conventional notion of wealth inequality that can be mapped into actual wealth inequality might help close the bridge with our simple tax structure approach.

[53] Condemning future generations with almost certainty to misery to improve the welfare and incentives of their ancestors runs diametrically against any modern notion of opportunity.

less than individual discounting or through a minimum bound on the continuation utility of any generation), the immiseration result disappears and an ergodic steady-state is obtained where each dynasty has the same long-run future expectations (as in our linear tax models). This illuminating result was established by Farhi and Werning, 2007 and Farhi and Werning 2010 show that it also applies in their NDPF bequest tax model.[54]

Farhi and Werning (2010) then show that instead of being zero in expectation as in Kocherlakota (2005), the optimal expected bequest tax rate is negative and given by essentially the same formula to the one obtained in their two-period model. This result has again an analog in our simple linear tax framework, comparing formula (14) from period 0 perspective to formula (16) from a steady-state perspective, we can see that the formula from the steady-state perspective adds a term in $\bar{b}^{\text{received}}$ that *reduces* τ_B (everything else being equal). Exactly as the Farhi-Werning vs. Kocherlakota case, this extra term arises because there is a welfare cost to bequest receivers of bequests taxation when we do not evaluate social welfare solely from period 0 perspective but also put some weight on later generations who are bequests receivers.

Hence, Farhi and Werning (2010), starting from two very different zero-tax benchmarks (Atkinson-Stiglitz in the 2-period model and Kocherlakota in the infinite horizon), and using the most general tax structures, obtain the same negative bequest tax deviation when including the welfare of future generations. In our view, however, the Kocherlakota zero-tax with immiseration benchmark is not meaningful for actual policy practice. Going from immiseration to an ergodic equilibrium with mobility is economically first order. The fact that the mobility ergodic equilibrium is achieved with negative bequest tax rates is second order. Something more fundamental much have changed from the Kocherlakota (2005) tax structure leading to immiseration to the Farhi and Werning (2010) tax structure leading to ergodic mobility.

Overall, we think that our "simple tax structure" analysis can be in large part reconciled with the approach of Farhi-Werning. We both agree that counting bequest receivers in social welfare leads to lower optimal bequest tax rates. Both approaches have merit and should be seen as complementary as they can cast useful light on each other.

It would be valuable to bridge the gap between our approach and the NDPF approach by considering intermediary tax systems more complex than our current linear tax structure yet not as general and history dependent as the NDPF general mechanism. In practice, modern democratic societies care about the rights and well-being of individuals–and not about the rights and well-being of dynasties–which generates strong horizontal equity constraints on tax structures.[55] At the micro-level, one individual's taxes should depend on her economic resources

[54] Our linear tax model always delivers an ergodic steady-state (either when maximizing period 0 utility or maximizing steady-state utility) and hence is not as critically sensitive to the social welfare criterion choice.

[55] As a practical example, it is worth noting that entails (namely the possibility for the funder of an estate to decide what future generations could do with the wealth) were abolished not only by the French Revolution but also by the American Revolution. As Thomas Jefferson once puts it: "the world belongs to the living" (see Beckert 2007, p.13).

(labor income and inherited wealth) but not on the history of how her inherited wealth was accumulated by her ancestors (or how her labor income abilities were transmitted to her). At the macro-level, future generations should not be collectively punished to improve the welfare and incentives of earlier generations, ruling out immiseration schemes, and requiring stationary tax systems.

Therefore in order to analyze how more complex tax mechanisms could in practice increase welfare, it would be valuable to restrict attention to tax schemes respecting some basic horizontal equity and stationarity properties. As compared to the simple linear and two-bracket tax structures analyzed in this paper, we see two main avenues for increasing welfare. First, and most obviously, one should look for the full non-linear optimum, and not just for two-bracket inheritance taxes. This introduces computational complications but does not radically change the optimal tax problem. Next, and maybe more importantly, one could use nonlinear taxes on joint current individual labor income and inherited wealth. I.e. how much one pays in inheritance taxes could also depend on how much one earns in terms of labor income. To our knowledge, such schemes have never been used in the real world. However it would be worth analyzing in future work to what extent they could generate substantial welfare improvements as compared to standard inheritance taxes (where the amount due is independent of one's labor income).

A.5 Calibration and Numerical Simulations Details

All detailed calibration results, computer codes and formulas are provided in the data appendix file available on line. Our main sensitivity checks are reported on Figures A1-A6 and are commented in section 4 of the paper. Many supplementary sensitivity checks are provided in the excel file. One can also use the file to change the parameters and graph the resulting optimal tax rates series, both for linear and two-bracket tax specifications (with thresholds at \$500,000 or € and \$1,000,000 or €). Here we clarify and highlight a number of technical issues and limitations of our calibrations, which should be better addressed in future research.

First, and most importantly, we did not try in our calibrations to correct for reporting biases in either EP 2010 or SCF 2010. This is potentially a serious problem, because respondents in wealth survey are known to massively underreport bequest and gift receipts. In France, the aggregate annual flow of bequests and gifts reported in household wealth surveys is less than 50% of the aggregate flow found in fiscal data - which is troubling, given that the latter ignores tax exempt assets such as life insurance, and hence is a lower bound for the true economic flow (see Piketty 2011).

In case the under-reporting rate is the same for all bequest receivers, then the distributional ratios $\bar{b}^{\text{received}}$ and $\bar{b}^{\text{left}}$ are unaffected, and our resulting optimal tax rates are unbiased.

However there are reasons to believe that reporting rates are not randomly distributed. For instance, it could be that individuals who have gone through a downward sloping wealth trajectory - i.e. who inherited 500 000$ twenty years ago and only have 100 000$ - tend to forget to report their inheritance more often than average. On the contrary, it could be that individuals with high current net worth like to present themselves as "self made" individuals and therefore tend to not to report bequests and gifts (even if they represent only part of their current wealth). It could also be that both types of under-reporting are present whenever bequest receipts are very large: large inheritors just tend to forget, whatever happens to their wealth trajectory.

Preliminary analysis of the data suggests that this latter biais is indeed what is happening, probably in both countries, and particularly so in the US: there are too few individuals reporting large bequests and gifts in the retrospective questionnaires (as compared to the number of decedents with large wealth in previous surveys). In both countries, a substantial fraction of the population actually reports no bequest or gift receipt at all. Per se this is not necessarily problematic: given the large concentration of wealth (bottom 50% receivers usually receive less than 5% of aggregate bequest flow), it is natural that the bottom half reports very little bequest and gift or not all; so what we did is to randomly attribute bequest received to bottom percentiles so as to obtain a continuous distribution and replicate the right wealth shares.[56] In France, about 50% of the population aged 70-year-old and over reports positive bequest or gifts (up from about 30% within the 18-to-29-year-old), which is consistent with tax data. In the US, however, it is only 30% (up from about 10% among the 18-to-29-year-old). This can be partly explained by the higher level of wealth inequality observed in the US, but this does not seem to be sufficient. Another possible explanation is the stigma associated to inheritance in US society (where "self made" values are particularly strong in moral and political discourses). Yet another possible explanation is the fact that the retrospective questionnaire is more detailed in the French wealth survey than in the US survey. In particular, the French survey asks separate questions about bequests and gifts received by each spouse, whereas there is only one question for both spouses in the SCF (so it is possible that the respondent sometime responds solely for himself or herself, although he or she is asked not to do so). In any case, there is a basic inconsistency between the self-reported bequest flow in current wealth survey and the theoretical bequest flow that one could compute by applying mortality rates to parental wealth reported in previous wealth surveys. This is likely to bias downwards optimal tax rates (if only a very small percentage of the population reports any positive bequest, then by construction zero receivers make the vast majority of the population and accumulate almost as much as the average, so that $\bar{b}^{\text{left}}$ is close to 100%, which leads to lower τ_B). This should be addressed in a systematic manner in future research.

We stress that some of the differences that we obtain between France and in the U.S. (in particular the fact that $\bar{b}^{\text{left}}$ within the bottom 50% receivers is as large as 70%-80% in the U.S., vs 60%-70% in France; see excel file) might well reflect such reporting biases, rather than

[56] We used a uniform law with upper bound equal to bottom reported bequests; we tried several specifications, and this made little difference to the resulting estimates. See excel file.

true differences in wealth mobility and hence socially optimal tax rates. The calibration results presented in this paper should be viewed as exploratory: they provide illustrative orders of magnitudes for key parameters and optimal tax rates, but should not be used to make fine policy recommendations or comparisons between countries.

In order to illuminate the crucial role played by wealth inequality and mobility, and the importance of using the right data sources to estimate these distributional parameters, we provide in the on-line appendix file detailed estimates using the micro files of estate tax returns collected by Piketty, Postel-Vinay and Rosenthal (2011) in the Paris archives over the 1872-1937 period. This is an interesting time period to look at, since it was characterized by large inheritance flows and extreme wealth concentration (with over 90% of aggregate inheritance received by top 10% successors). In addition, this is highly reliable, exhaustive administrative data covering wealth over two generations (something that is usually difficult to do), which does not suffer from the same self-reporting biases as the contemporary survey data. We find that $\bar{b}^{\text{left}}$ is as low as 20%-30% for the bottom 80% receivers (maybe with a slight rise over the period). This would imply very high optimal inheritance tax rates - typically above 80% for the benchmark values parameters used here.[57] This would also suggest that wealth mobility has increased quite spectacularly between Paris 1872-1937 and either France 2010 or the US 2010 (which would make sense, given the decline in both the aggregate level of inheritance flows and the concentration of inherited wealth). However given the data sources biases for the recent period, it is difficult to make a precise comparison. It would be valuable to use similar administrative data for the recent period. We leave this to future research.

Next, it would be valuable to introduce individual specific estimates for the strength of bequest motive ν (using available questionnaires) and for capitalization factors (here we applied the same annual real rate of return to all bequests and gifts; this seems to have rather limited impact on optimal tax rates, however; see excel file).

Next, it would be interesting to use our estimates to compute the full social optimum implied by various social welfare functions, in particular the utilitarian optimum. In effect, this would amount to computing a weighted average of the optimal tax rates depicted on Figure 1, with weights given by the marginal social value of extra income for the different percentiles of the distribution of bequest received. The exact result will depend with the curvature γ, but it is pretty obvious that for any reasonably large curvature (putting sufficiently more weights on bottom deciles), the utilitarian optimum will be very close to the bottom 70% receivers' most preferred tax rate. A more complicated issue is to decide whether one should use the same curvature within each percentile of the distribution of bequest received. In effect, our calibrations ignore redistribution issues between individuals in the same percentile of bequest received, but with different labor incomes. The full social welfare optimum should also introduce

[57] Note also that it is possible that the $\bar{y}_L$ effect pushes in the same direction: in a rentier society where the very rich do not work, then $\bar{y}_L$ can be larger than 100% for the poor and the middle class. Unfortunately we do not observe labor earnings in estate tax returns, so we cannot really say.

this dimension of redistribution.

Finally, it would be valuable to introduce more structure into our calibrations. In our baseline estimates, we simply compute the optimal tax rates by plugging observed distributional ratios into the optimal tax formula. However in practice distributional ratios should respond to change in tax rates, thereby implying that our baseline estimates are biased upwards. In particular, one needs to put a minimum structure so that $\bar{b}^{\text{left}}$ depends on τ_B. In the case $\tau_B = 100\%$, $\bar{b}^{\text{left}} = \bar{y}_L$ is natural (as zero receivers are no longer disadvantaged). The simplest way to proceed is to consider that we estimate $\bar{b}^{\text{left}}$ at the current rate τ_B^{current}, and then assume that $\bar{b}^{\text{left}}(\tau_B)$ is linear in τ_B (this is what we obtain in the linear savings model, see Piketty and Saez 2012):

$$\bar{b}^{\text{left}}(\tau_B) = \frac{\bar{b}^{\text{left}}(\tau_B^{\text{current}})(1-\tau_B) + (\tau_B - \tau_B^{\text{current}})\bar{y}_L}{1-\tau_B^{\text{current}}}$$

The main difficulty with this approach is that one needs to specify the current tax system, which in practice is highly non-linear, and relies much more on the annual taxation of the flow of capital income and corporate profits (and on annual property or wealth taxes) and on inheritance taxes. Taking all forms of capital taxes together, the average effective capital tax rate is about 30%-40% in both France and the US. Preliminary estimates using this simplified view of the current tax system lead to the conclusion that the extra effects implied by the linear structure would not be very large - as long as the optimal tax rate is not too different from the current one. For instance, if we take τ_B^{current}=40%, and if we start from a situation where $\tau_B = 60\%$, (which is approximately the optimal linear inheritance tax rate for bottom 70% receivers in both France and the US, see Figure 1), then the new corrected optimal tax rate would be reduced to $\tau_B \simeq 55\%$. We leave more sophisticated calibrations - in particular taking into account the non-linear structure of the tax system - to future research.

Another major limitation of our calibrations is that we compute optimal tax rates from the viewpoint of single cohort, namely individuals over 70-year-old in 2010. This corresponds to the cohorts born in the 1920s-1930s, who for the most part received bequests from their parents in the 1970s-1980s, and who are about to leave bequests to their children in the 2010s-2020s. The problem is that we are not in a steady-state. In France, the aggregate annual flow of bequest was slightly over 5% of national income in the 1970s, and has gradually increased in recent decades, up to about 15% of national income in the 2010s (Piketty, 2011); in the U.S., the trend is going in the same direction, though probably with a lower slope.[58] In other words, we have computed optimal tax rates from the viewpoint of cohorts who at the aggregate level have received less bequests than what they will leave - which biases downwards optimal rates.

[58]See the discussion in Piketty 2011. See also the series by Piketty and Zucman 2012 showing that the aggregate wealth-income ratio has increased significantly in the US since the 1970s, but less strongly than in Europe. This, together with a stronger demography (younger population and lower mortality rates) and larger non-transmissible, annuitized wealth (pension funds), is likely to deliver a smaller rise in the aggregate bequest flow.

In the working paper version of this work (Piketty and Saez 2012), we show that the optimal tax formula can be re-expressed in terms of the aggregate bequest flow $b_y = B/Y$, and we present calibrations illustrating the fact that for a given structure of preferences and shocks, the optimal tax rate is a steeply increasing function of b_y. The intuition is the following: with a low b_y, there is not much gain from taxing high bequest receivers from my own cohort, and in addition low and high bequest receivers accumulate wealth levels that are not too far apart. In future research, it would be valuable to combine the micro calibrations emphasized here and the macro calibrations presented in the working paper in order to compute cohort-varying, out-of-steady-state optimal tax rates. It is likely that the optimal tax rates from the viewpoint of more recent cohorts will be significantly larger than those for older cohorts.

A.6 Extensions not Covered in the Main Text

A.6.1 Optimal Nonlinear Inheritance Taxation

Our formulas can be extended to the case with nonlinear bequest taxation when the nonlinear bequest tax takes the following simple but realistic form. Bequests below a threshold b_t^* are exempt and the portion of bequests above the threshold b_t^* is taxed at the constant marginal tax rate τ_{Bt}. In effect the tax on b_{ti} is $\tau_{Bt}(b_{ti} - b_t^*)^+$. Actual bequest tax systems often do take such a form. Considering multiple brackets with different rates is unfortunately intractable as we explain below. We consider only the basic model of Section 2 and the Meritocratic Rawlsian criterion (the formulas can also be extended to the models of Section 3 as well). We consider the case with "generational" budget balance so as to be consistent with dynamic efficiency (as is possible when considering the zero-receivers optimum as discussed in appendix A.3).

Let us denote by $B_{ti} = (b_{ti} - b_t^*)^+$ taxable bequests of individual ti and $B_t = \int_i B_{ti}$ aggregate taxable bequests.

The individual maximization problem is:

$$\max_{c_{ti}, b_{t+1i} \geq 0} V^{ti}(c_{ti}, R[b_{t+1i} - \tau_{Bt+1}(b_{t+1i} - b_{t+1}^*)^+], l_{ti}) \text{ s.t. } c_{ti} + b_{t+1i} = R[b_{ti} - \tau_{Bt} B_{ti}] + w_{ti} l_{ti}(1 - \tau_{Lt}) + E_t.$$

The individual first order condition for bequests left is $V_c^{ti} = R(1 - \tau_{Bt+1}) V_{\underline{b}}^{ti}$ if $B_{t+1i} > 0$ and $V_c^{ti} = R V_{\underline{b}}^{ti}$ if $0 < b_{t+1i} < b_{t+1}^*$. Importantly, $B_{t+1i} V_c^{ti} = R(1 - \tau_{Bt+1}) B_{t+1i} V_{\underline{b}}^{ti}$ is always true.

We take b^* as given and constant with t in the steady state. The government solves

$$SWF = \max_{\tau_L, \tau_B} \int_i \omega_{ti} V^{ti}(R(b_{ti} - \tau_B B_{ti}) + w_{ti} l_{ti}(1 - \tau_L) + E_t - b_{t+1i}, R(b_{t+1i} - \tau_B B_{t+1i}), l_{ti}). \quad \text{(A13)}$$

with E given and τ_L and τ_B linked to meet the "generational" budget constraint, $E = \tau_B B_{t+1} + \tau_L y_{Lt}$. The aggregate variable B_{t+1} is a function of $1 - \tau_B$ (assuming that τ_L adjusts), and y_{Lt} is a function of $1 - \tau_L$ (assuming that τ_B adjusts). Formally, we can define the corresponding long-run elasticities as:

$$e_B = \frac{1 - \tau_B}{B_t} \frac{dB_t}{d(1 - \tau_B)}\bigg|_E \quad \text{and} \quad e_L = \frac{1 - \tau_L}{y_{Lt}} \frac{dy_{Lt}}{d(1 - \tau_L)}\bigg|_E.$$

Consider a small reform $d\tau_B > 0$, budget balance with $dE = 0$ requires that $d\tau_L$ is such that:

$$B_{t+1} d\tau_B \left(1 - e_B \frac{\tau_B}{1-\tau_B}\right) = -d\tau_L y_{Lt} \left(1 - e_L \frac{\tau_L}{1-\tau_L}\right).$$

Using the fact that b_{t+1i} and l_{ti} are chosen to maximize individual utility and applying the envelope theorem, the fact that $R(b_{ti} - \tau_B B_{ti}) \equiv 0$ for zero-receivers, the effect of the reform $d\tau_B, d\tau_L$ on steady-state social welfare is:

$$dSWF = \int_i \omega_{ti} V_c^{ti} \cdot (-d\tau_L y_{Lti}) + \omega_{ti} V_{\underline{b}}^{ti} \cdot (-d\tau_B R B_{t+1i}).$$

At the optimum, $dSWF = 0$. Using the individual first order condition $V_c^{ti} B_{t+1i} = R(1 - \tau_B) B_{t+1i} V_{\underline{b}}^{ti}$, and the expression above for $d\tau_L$, and the definition of g_{ti}, we have:

$$0 = \int_i g_{ti} \cdot \left[\frac{1 - \frac{e_B \tau_B}{1-\tau_B}}{1 - \frac{e_L \tau_L}{1-\tau_L}} \frac{y_{Lti}}{y_{Lt}} B_{t+1} d\tau_B - \frac{d\tau_B B_{t+1i}}{1-\tau_B} \right],$$

Let $\bar{y}_L$, $\bar{B}^{\text{left}}$ be the population averages of $g_{ti} \cdot y_{Lti}/y_{Lt}$, $g_{ti} \cdot B_{t+1i}/B_{t+1}$. Dividing by $B_{t+1} d\tau_B$, the first order condition is rewritten as:

$$0 = \frac{1 - e_B \tau_B/(1-\tau_B)}{1 - e_L \tau_L/(1-\tau_L)} \bar{y}_L - \frac{\bar{B}^{\text{left}}}{1-\tau_B}.$$

Finally, as in optimal top labor income taxation (Saez, 2001), we can define the elasticity e_b of top bequests (i.e., the full bequests among taxable bequests) with respect to $1 - \tau_B$. This elasticity e_b is typically the one estimated in empirical studies (e.g., Kopczuk and Slemrod, 2001). It is related to elasticity of aggregate taxable bequests e_B through the Pareto parameter a of the bequests distribution through the simple equation $e_B = a \cdot e_b$ with $a = b^m(b^*)/[b^m(b^*) - b^*]$ where $b^m(b^*)$ is the average bequest among bequests above the taxable threshold b^*. To see this, note that for taxable bequests, $b_{ti} - b^* = B_{ti}$ so that $b_{ti} \frac{db_{ti}}{b_{ti}} = (b_{ti} - b^*) \frac{dB_{ti}}{B_{ti}}$, and hence $b_{ti} e_{bti} = (b_{ti} - b^*) e_{Bti}$ at the individual level. Aggregating across all taxable bequests, we get $b^m(b^*) e_b = (b^m(b^*) - b^*) e_B$, i.e., $a \cdot e_b = e_B$. Hence, we can state:

Nonlinear Top Rate Steady-State Rawlsian Meritocratic Optimum. *The optimal tax rate τ_B above threshold b^* that maximizes long-run steady state social welfare of zero-receivers with "generational" budget balance is given by:*

$$\tau_B = \frac{1 - \left[1 - \frac{e_L \tau_L}{1-\tau_L}\right] \cdot \frac{\bar{B}^{\text{left}}}{\bar{y}_L}}{1 + e_B} = \frac{1 - \left[1 - \frac{e_L \tau_L}{1-\tau_L}\right] \cdot \frac{\bar{B}^{\text{left}}}{\bar{y}_L}}{1 + a \cdot e_b}. \tag{A14}$$

where $\bar{B}^{\text{left}}$ and $\bar{y}_L$ are the average taxable bequests and average labor income among zero-receivers (relative to population wide averages), e_B is the elasticity of aggregate taxable bequests, a is the Pareto parameter of the bequest distribution, and e_b the elasticity of full bequests (among taxable bequests).

Three points are worth noting. First, if zero-receivers never accumulate a bequest large enough to be taxable, then $\bar{B}^{\text{left}} = 0$, and the formula reverts to the revenue maximizing tax rate $\tau_B = 1/(1+e_B) = 1/(1+a \cdot e_b)$.[59] Second, as b^* grows, there are two possibilities: either $\bar{B}^{\text{left}}$ converges to zero or converges to a positive level. The first case corresponds to an aristocratic society where top bequests always come from past inheritances and never solely from self-made wealth. In that case again, the optimum τ_B should be the revenue maximizing rate. The second case corresponds to a partly meritocratic society where some of the top fortunes are self-made. In that case, even for very large b^*, zero-receivers want a tax rate on bequests strictly lower than the revenue maximizing rate. In reality, it is probable that $\bar{B}^{\text{left}}$ declines with b^* as the fraction of self-made bequests left likely declines with the size of bequests. If the elasticity e_b and a are constant, then this suggests that the optimum τ_B increases with b^*. The countervailing force is that aristocratic wealth is more elastic as the bequest tax hits those fortunes several times across several generations, implying that e_b might actually grow with b^*.[60]

Third, real world estate tax systems generally have several progressive rates, and ideally one would like to solve for the full non-linear optimum. Unfortunately there is no simple formula for the optimal nonlinear bequest tax schedule. The key difficulty is that a change in the tax rate in any bracket will end up having effects throughout the distribution of bequests in the long-run ergodic equilibrium. This difficulty does not arise in the simple case where there is a single taxable bracket and we consider the zero-receiver optimum. We leave further exploration of full non linear optima to future research.

A.6.2 Equivalence between Consumption and Labor Tax

Let us denote by τ_{Ct} the tax-inclusive consumption tax rate, i.e., an individual spending \$1 nominal dollar in period t, actually consumes $\$(1-\tau_{Ct})$ in real terms and pays τ_{Ct} in taxes with a consumption tax at rate τ_{Ct}.[61] When there is a consumption tax, let us denote by c^C_{ti}, b^C_{ti}, E^C_{ti} the nominal value of consumption, bequests received, and demogrant and by $c_{ti} = (1-\tau_{Ct})c^C_{ti}$, $b_{ti} = (1-\tau_{Ct})b^C_{ti}$, $E_t = (1-\tau_{Ct})E^C_t$, the real levels (in term of purchasing power) of consumption, bequests received, and demogrant taking into account the consumption tax.

We can prove the following general equivalence between a labor tax and a consumption tax when the government can use debt.

Proposition 1 ***Equivalence between labor tax and consumption tax.*** *Consider a labor tax, bequest tax system $(\tau_{Lt}, \tau_{Bt})_{t \geq 0}$ (with zero consumption tax) with aggregate consumption c_t, aggregate private bequests b_t, government net assets a_t, demogrant E_t, and capital $K_t = a_t + b_t$ (as in Section 3.1 in the text).*

[59] The formula takes the same form as in standard optimal labor income tax theory (see Saez 2001).

[60] This is easily seen in the model with binomial random tastes (see Piketty and Saez, 2012).

[61] Such an tax-inclusive consumption tax τ_C is equivalent to a conventional tax-exclusive consumption tax at rate t_c if $1-\tau_C = 1/(1+t_c)$. We use the tax-inclusive τ_C concept to obtain a direct equivalence with τ_L (as labor income taxes are conventionally expressed in tax-inclusive terms).

• *This system is economically exactly equivalent to a consumption tax, bequest tax system* $(\tau_{Ct} = \tau_{Lt}, \tau_{Bt})_{t\geq 0}$ *(with zero labor tax) along with additional subsidies/taxes for bequests at rate* $s_t = (\tau_{Ct+1} - \tau_{Ct})/(1-\tau_{Ct})$*, with aggregate nominal consumption* $c_t^C = c_t/(1-\tau_{Ct})$*, aggregate nominal bequests* $b_t^C = b_t/(1-\tau_{Ct})$*, government net assets* $a_t^C = a_t - \tau_{Ct} b_t^C$*, nominal demogrant* $E_t^C = E_t/(1-\tau_{Ct})$*, and capital* $K_t = a_t^C + b_t^C = a_t + b_t$*.*
• *If initial labor taxes are time invariant, the equivalent consumption tax is also time invariant and no additional bequest subsidies/taxes are needed* ($s_t = 0$)*.*
• *It is possible to transition from the labor tax to the consumption tax system in (any) period* t_0 *with no change in any individual utilities nor real variables by issuing debt* $\tau_{Ct_0} b_{t_0}/(1-\tau_{Ct_0})$ *at the end of period* $t_0 - 1$*. The difference in subsequent permanent tax yield exactly pays the interest on this extra-debt.*

Proof:
Individual equivalence. Under the consumption tax $(\tau_{Ct}, \tau_{Bt})_{t\geq 0}$, the individual utility is:

$$V^{ti}(c_{ti}, (1-\tau_{Bt+1})R_{t+1}b_{t+1i}, l_{ti}) = V^{ti}((1-\tau_{Ct})c_{ti}^C, (1-\tau_{Ct+1})(1-\tau_{Bt+1})R_{t+1}b_{t+1i}^C, l_{ti}),$$

i.e., we assume that individuals care about the real value of the bequests as wealth matters in real terms, not nominal terms. With a subsidy at rate s_t on bequests left b_{t+1i}^C, the individual budget constraint in nominal terms is

$$c_{ti}^C + (1-s_t)b_{t+1i}^C \leq y_{Lti} + (1-\tau_{Bt})R_t b_{ti}^C + E_t^C, \tag{A15}$$

so that multiplying by $1-\tau_{Ct}$ and using real quantities, we have:

$$c_{ti} + (1-s_t)\frac{1-\tau_{Ct}}{1-\tau_{Ct+1}} b_{t+1i} \leq (1-\tau_{Ct})y_{Lti} + (1-\tau_{Bt})R_t b_{ti} + E_t.$$

Hence, from the individual point of view, the consumption tax system with the subsidy $(\tau_{Ct} = \tau_{Lt}, \tau_{Bt}, s_t)_{t\geq 0}$ is equivalent to the standard labor tax $(\tau_{Lt} = \tau_{Ct}, \tau_{Bt})_{t\geq 0}$ (with zero consumption tax) when the subsidy s_t is set so that $(1-s_t)\frac{1-\tau_{Ct}}{1-\tau_{Ct+1}} = 1$, i.e., $s_t = (\tau_{Ct+1} - \tau_{Ct})/(1-\tau_{Ct})$ as in the Proposition. The subsidy (or tax) s_t is needed to undo the inter-temporal distortion created by time varying consumption taxes (as time-varying labor taxes do not create such inter-temporal distortions). If the consumption tax rate increases, $s_t > 0$ and conversely if the consumption tax rate decreases, $s_t < 0$. Naturally, $s_t = 0$ if the consumption tax rate does not vary from t to $t+1$.

Government and macro-level equivalence. Aggregating the individual budget constraints (A15), we have:

$$c_t^C = y_{Lt} + (1-\tau_{Bt})R_t b_t^C + E_t^C - (1-s_t)b_{t+1}^C. \tag{A16}$$

Let us denote by a_t^C the government net assets in period t in this consumption tax scenario. Government assets a_t^C and private assets b_t^C form the capital stock so that we have $K_t = b_t^C + a_t^C$.

Those assets are invested and earn a rate of return R_t at the end of period t (R_t is exogenous in the open economy case and endogenous and equal to $R_t = 1 + F_K(K_t, L_t)$ in the closed economy case). At the end of period t, the government collects period t taxes, provides the subsidy s_t to bequests left to period $t+1$, pays the demogrant E_t^C. Hence, the dynamic government budget constraint is:

$$a_{t+1}^C = R_t a_t^C + \tau_{Bt} R_t b_t^C + \tau_{Ct} c_t^C - s_t b_{t+1}^C - E_t^C,$$

Using the expression (A16) for c_t^C, we have

$$a_{t+1}^C = R_t a_t^C + \tau_{Bt} R_t b_t^C + \tau_{Ct}[y_{Lt} + (1-\tau_{Bt}) R_t b_t^C + E_t^C - (1-s_t) b_{t+1}^C] - s_t b_{t+1}^C - E_t^C,$$

and hence,

$$a_{t+1}^C + [s_t + \tau_{Ct}(1-s_t)] b_{t+1}^C = R_t[a_t^C + \tau_{Ct} b_t^C] + \tau_{Ct} y_{Lt} + \tau_{Bt}(1-\tau_{Ct}) R_t b_t^C - (1-\tau_{Ct}) E_t^C.$$

Denoting $a_t = a_t^C + \tau_{Ct} b_t^C$ and using the fact that $s_t + \tau_{Ct}(1-s_t) = \tau_{Ct+1}$, we have

$$K_t = a_t^C + b_t^C = a_t + (1-\tau_{Ct}) b_t^C = a_t + b_t, \text{ and}$$

$$a_{t+1} = R_t a_t + \tau_{Ct} y_{Lt} + \tau_{Bt} R_t b_t - E_t.$$

Those two equations are the capital stock equation and the government dynamic budget equation of the labor tax system $(\tau_{Lt} = \tau_{Ct}, \tau_{Bt})_{t\geq 0}$. Hence, at the macro-economic level, the consumption tax with bequest tax and subsidy $(\tau_{Ct}, \tau_{Bt}, s_t)_{t\geq 0}$ is also equivalent to the labor tax with bequest tax $(\tau_{Lt} = \tau_{Ct}, \tau_{Bt})_{t\geq 0}$. This proves the first point in the proposition. The second point follows immediately as $s_t = 0$ when $\tau_{Ct} = \tau_{Ct+1}$.

Transition from labor tax to consumption tax. To understand this result and the role of initial conditions, it is pedagogically useful to consider a transition from the labor tax system $(\tau_{Lt}, \tau_{Bt})_{t\geq 0}$ to the consumption tax system (with bequest subsidy) $(\tau_{Ct} = \tau_{Lt}, \tau_{Bt}, s_t)_{t\geq 0}$. Suppose the transition happens in period t_0. In the old regime, bequests and government assets would have been b_{t_0} and a_{t_0}. In the new regime, bequests and government assets will be $b_{t_0}^C$ and $a_{t_0}^C$.

Period $t_0 - 1$ bequest leavers realize that their heirs will pay supplementary consumption tax in period t_0. Hence, the government needs to subsidize their bequests to undo the effect of the newly enacted τ_{Ct_0} with a subsidy at rate $s_{t_0} = \frac{\tau_{Ct_0}}{1-\tau_{Ct_0}}$. With this subsidy in place, period $t_0 - 1$ bequest leavers are unaffected by the new consumption tax and leave a new bigger nominal bequests $b_{t_0}^C = b_{t_0}/(1-\tau_{Ct_0})$ (relative to b_{t_0} they would have left in old labor tax regime). This requires the government to issue new debt $a_{t_0}^C - a_{t_0} = -b_{t_0}^C \tau_{Ct_0} = -b_{t_0} \frac{\tau_{Ct_0}}{1-\tau_{Ct_0}}$. Period t_0 capital is $K_{t_0} = a_{t_0}^C + b_{t_0}^C = a_{t_0} - b_{t_0}^C \tau_{Ct_0} + b_{t_0}^C = a_{t_0} + b_{t_0}$ i.e., unchanged relative to the capital stock in the old regime. Effectively, the new debt is fully absorbed by the new bigger bequests left. As we have seen, period t_0 generation faces exactly the same real budget constraints and hence makes the same real choices. We can re-derive the calculation above showing complete equivalence between the new government budget constraint and what the old one would have been:

$$a^C_{t_0+1} = R_{t_0} a^C_{t_0} + \tau_{Bt_0} R_{t_0} b^C_{t_0} + \tau_{Ct_0 t} c^C_{t_0} - s_{t_0} b^C_{t_0+1} - E^C_{t_0} \Leftrightarrow a_{t_0+1} = R_{t_0} a_{t_0} + \tau_{Ct_0} y_{Lt_0} + \tau_{Bt_0} R_{t_0} b_{t_0} - E_{t_0}.$$

This implies in particular that the new tax system raises the same revenue as the old system plus the interest to service the new extra-debt. Q.E.D.

Three points are worth noting about this proposition and proof.

First and most important, our optimal tax formulas for τ_B carry over unchanged in the case with consumption taxation (by just replacing τ_L by τ_C) in all the cases where we allow for government debt. As discussed in the main text, we think that the case with debt is the most appealing so that we can disconnect dynamic efficiency issues from cross-sectional redistribution issues. Intuitively, a consumption tax (relative to a labor tax) looks like it rewards savers and penalizes spenders, but in reality, with a labor tax, all bequests fall so that savers do not need to save as much and inheritors receive less. Hence, the situation is actually equivalent for everybody. Of course, a new consumption tax also taxes initial accumulated wealth. However, if the government compensates initial wealth holders for the new consumption tax, in effect, the consumption tax is a labor tax and none of the macro-economic variables are affected. If the government does not adjust debt, a shift from labor to consumption will tax initial wealth and hence will create real changes.[62]

Second, in the OLG case with 2-period long lives and with labor income only in period 1 (as in Section 3.2), the equivalence is even simpler because there is never any need to use s_t. If we consider a generation specific consumption tax τ^t_C that applies to the cohort young in period t and old in period $t+1$. Then, $(\tau_{Lt}, \tau_{Bt})_{t \geq 0}$ is equivalent to $(\tau^t_C = \tau_{Lt}, \tau_{Bt})_{t \geq 0}$ with no need of savings subsidies (as the tax rate on consumption is constant within a lifetime). Again, switching from labor to consumption tax in period t requires compensating the old for the new consumption tax (as they have already paid the labor tax) and the full micro- and macro-equivalence is preserved.

Third, the equivalence between labor and consumption taxes breaks down when using progressive taxes as the variation over time in the progressive consumption tax cannot be replicated with a time varying labor tax. A progressive consumption tax can therefore target idle heirs better than a progressive labor tax.

This general equivalence carries over more generally to a situation with annual capital income taxes. In that context, a labor tax is again equivalent to a consumption tax cum capital income subsidy s_t to undo the inter-temporal distortion created by time varying consumption taxes. Concretely, whenever the consumption tax varies over the time, the government has to compensate wealth holders period by period to keep the purchasing power of wealth constant. This can be done by issuing debt with no macro-economic consequences and the consumption tax is fully equivalent to a labor tax.

[62] Piketty and Saez (2012) discuss those effects.

Additional Appendix References

Beckert, Jens 2007. *Inherited Wealth*, Princeton University Press.

Farhi, Emmanuel and Iván Werning 2007. "Inequality and Social Discounting", *Journal of Political Economy*, 115(3), 365-402.

Kaplow, Louis 2006. "On the Undesirability of Commodity Taxation Even When Income Taxation Is Not Optimal." *Journal of Public Economics*, 90(6-7): 1235-50.

Kocherlakota, Narayana 2005. "Zero Expected Wealth Taxes: A Mirrlees Approach to Dynamic Optimal Taxation", *Econometrica*, 73, 1587–1621.

Laroque, Guy R. 2005. "Indirect Taxation is Superfluous under Separability and Taste Homogeneity: A Simple Proof." *Economics Letters*, 87(1): 141-4.

Saez, Emmanuel 2001. "Using Elasticities to Derive Optimal Tax Rates", *Review of Economic Studies*, 68, 205-229.

Figure A1: Optimal linear inheritance tax rates, by percentile of bequest received (France, variants with diff. eb = long-run bequest elasticity)

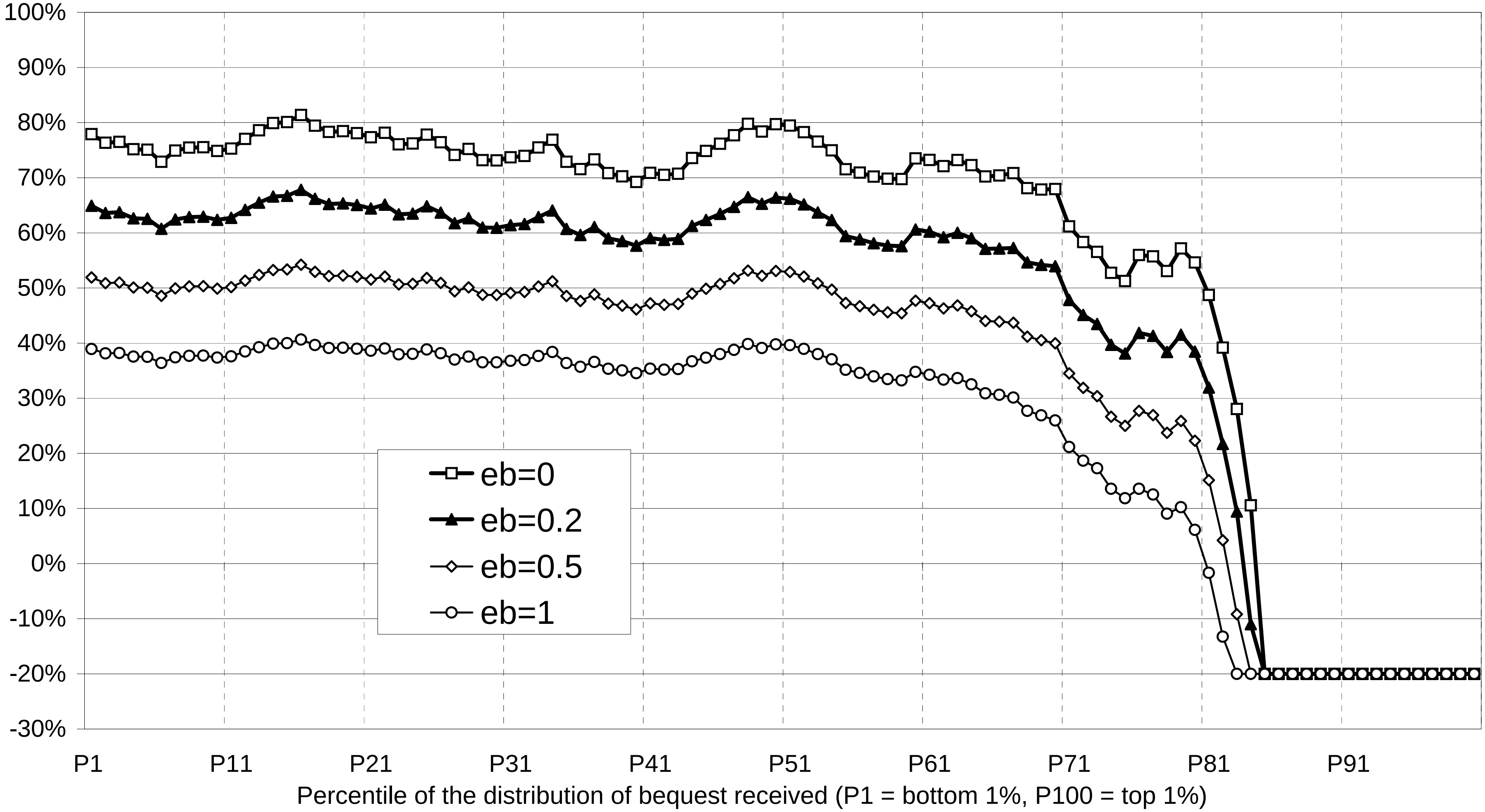

Figure A2: Optimal linear inheritance tax rates, by percentile of bequest received (U.S., variants with diff. eb = long-run bequest elasticity)

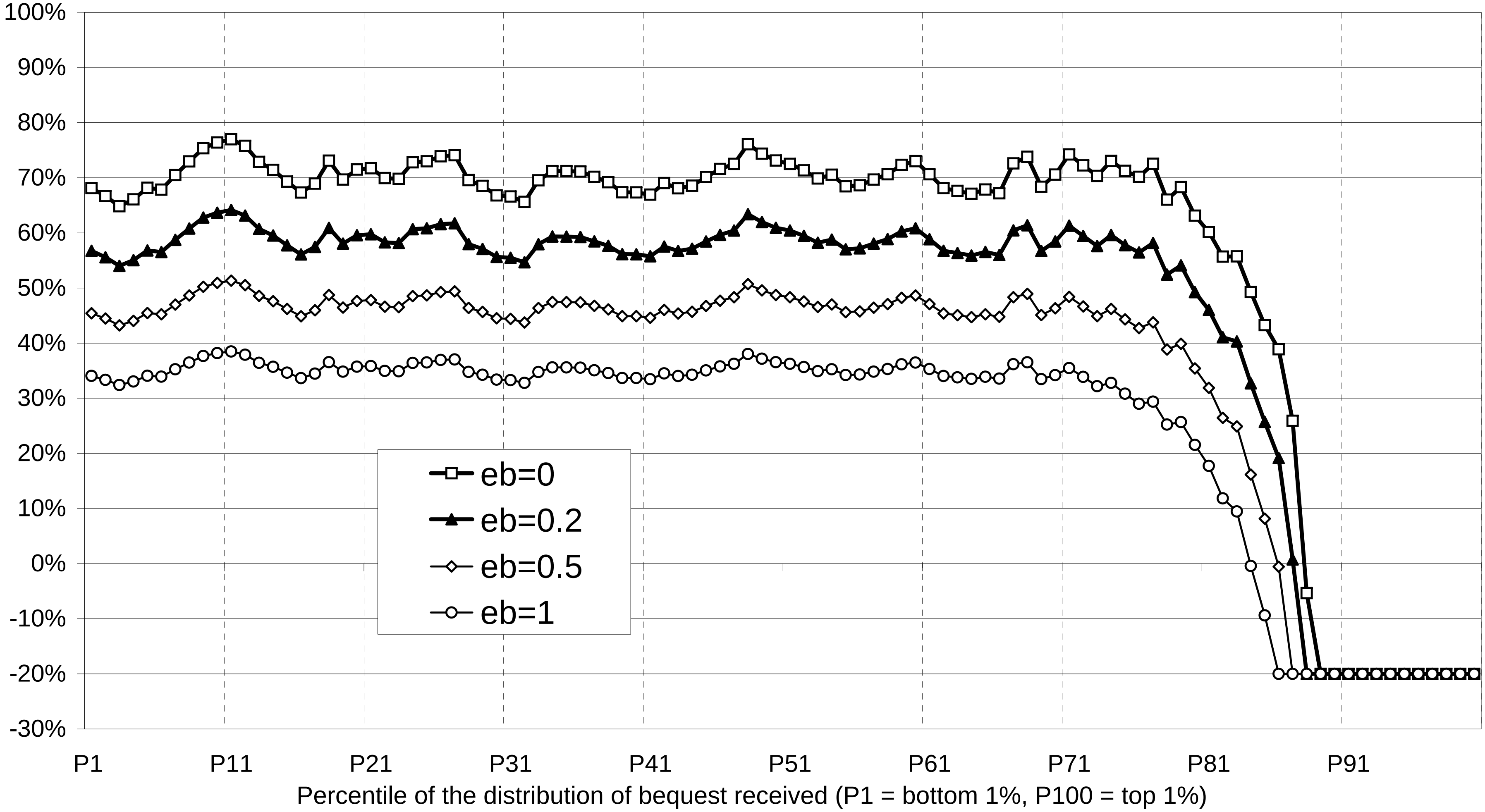

Figure A3: Optimal linear inheritance tax rates, by percentile of bequest received (France, variants with diff. v = strength of bequest motive)

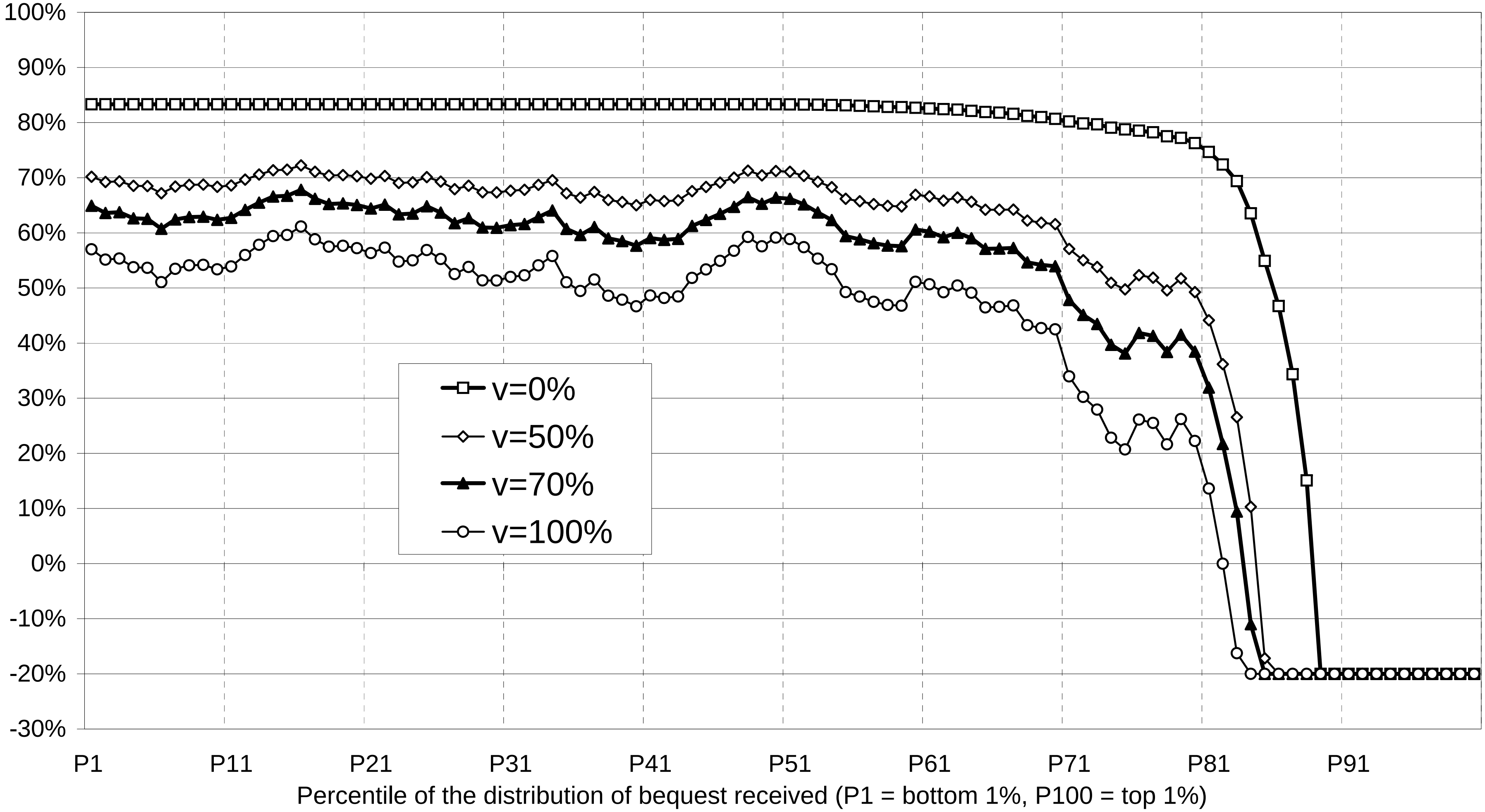

Figure A4: Optimal linear inheritance tax rates, by percentile of bequest received (U.S., variants with diff. v = strength of bequest motive)

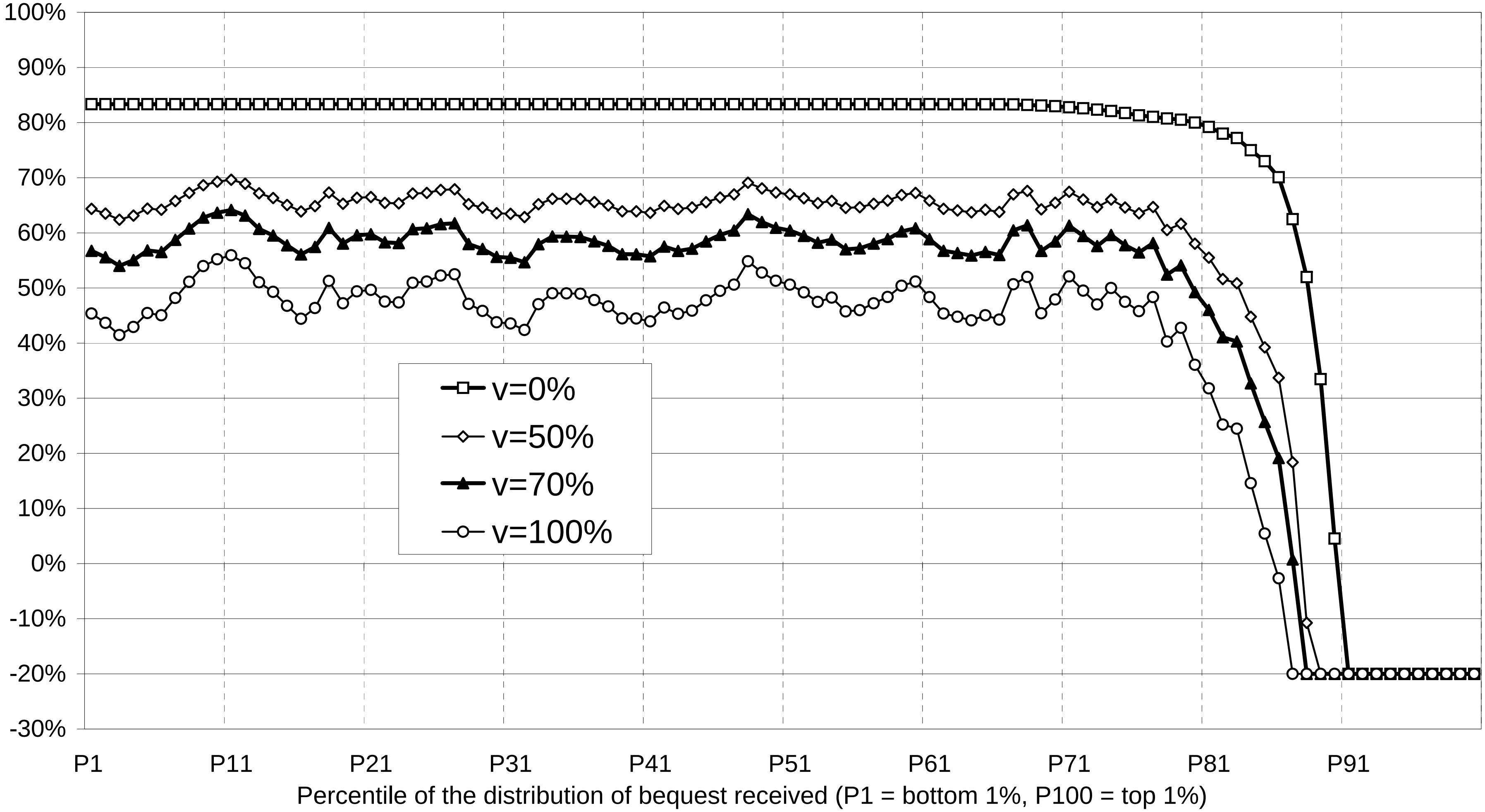

Figure A5: Optimal linear inheritance tax rates, by percentile of bequest received (France, variants with diff. r-g = capitalization factor)

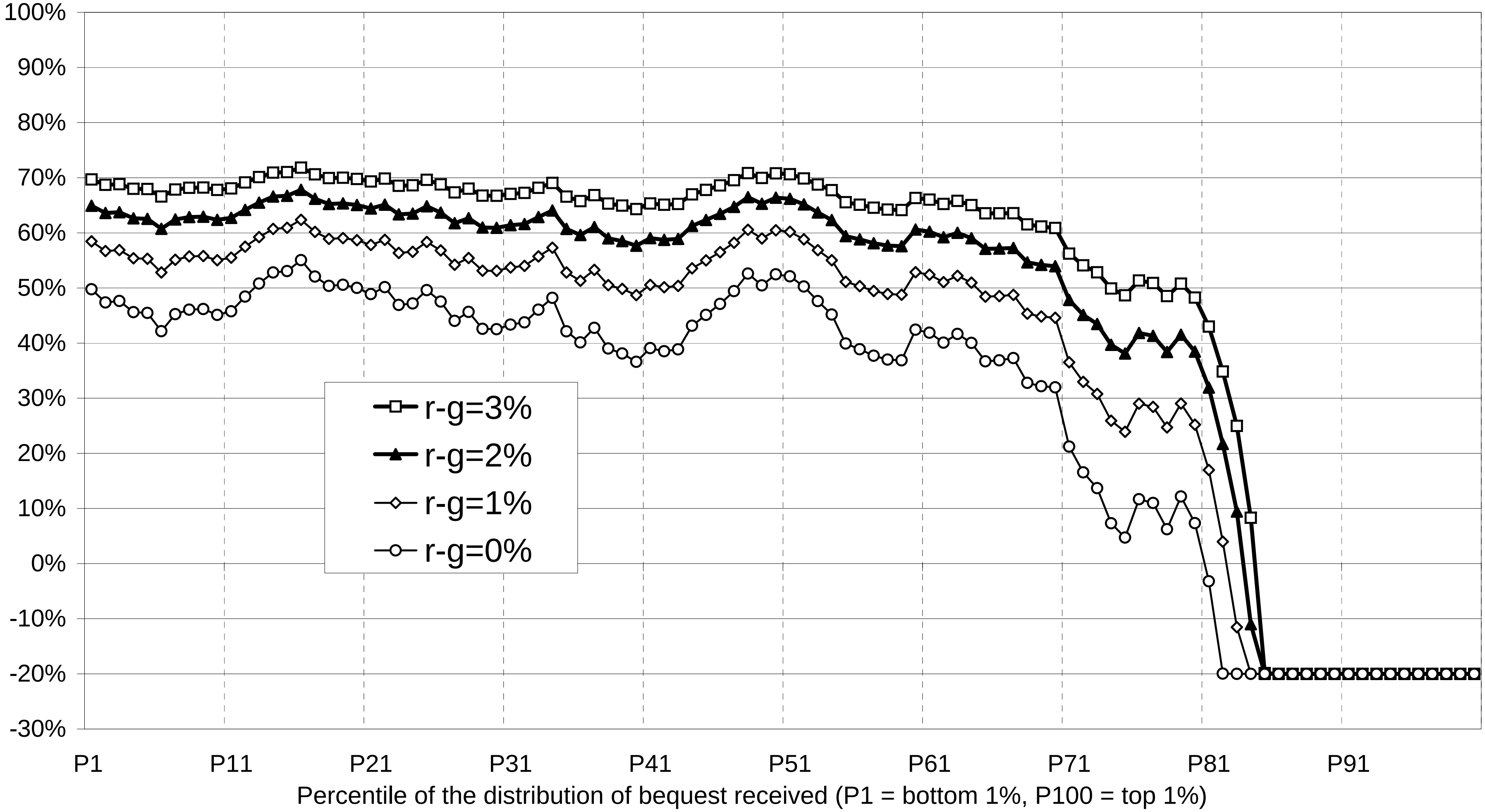

Figure A6: Optimal linear inheritance tax rates, by percentile of bequest received (U.S., variants with diff. r-g = capitalization factor)