Bavarian Graduate Program in Economics

BGPE Discussion Paper

No. 217

Optimal Estate Taxation: More (about) Heterogeneity across Dynasties

Philipp Krug

May 2022

ISSN 1863-5733

Editor: Prof. Regina T. Riphahn, Ph.D.
Friedrich-Alexander-Universität Erlangen-Nürnberg
© Philipp Krug

BGPE Discussion Paper Series

Optimal Estate Taxation: More (about) Heterogeneity across Dynasties

Philipp Krug*

December 7, 2021

Abstract

Standard models on optimal estate taxation do not allow for intergenerational transmission of bequest motives. However, correlation in bequest motives may exist due to genetic and cultural transmission of preferences or indirect reciprocity. I introduce such intergenerational correlation to a simple model with heterogeneously altruistic parents. I derive two insights for optimal linear estate taxation under a Utilitarian welfare measure. First, this correlation implies a higher optimal estate tax rate. Second, estate tax rates should be higher for those parents who inherited themselves.

Keywords: estate taxation, inheritance taxation, indirect reciprocity, intergenerational preference transmission

JEL Classification: H21, H24, D64

---

*KU Eichstätt-Ingolstadt, Ingolstadt, Germany (pkrug@ku.de). I would like to thank Alan Auerbach, Dominika Langenmayr, Dominik Sachs, Samina Sultan and participants at the IIPF Annual Congress in Tampere, the ZEW Public Finance Conference, the RSIT conference in Tübingen, and at seminars in Ingolstadt for helpful comments. Any remaining errors are my own. I worked on this project while visiting UC Berkeley; I gratefully acknowledge their hospitality as well as financial support by the Bavarian Graduate Program in Economics.

1 Introduction

Wealth transfers (i.e., gifts and bequests) are one important determinant of wealth accumulation and lifetime inequality in the long run (Piketty, 2011). Wealth transfers can explain approximately half of the wealth correlation across two generations and are an important determinant of wealth persistence even across three generations (Adermon *et al.*, 2018). This may be partly explained by the persistence of bequest motives across generations. While there is agreement that the optimal tax design crucially depends on the individuals' bequest motives (see Kopczuk, 2013b), little has been said about how intergenerational correlation of bequest motives affects optimal estate taxation.

Bequest motives may be correlated across generations for at least three reasons. A first channel is the existence of retrospective bequest motives. The roots of retrospective bequest motives lie in 'indirect reciprocity' (first named by Alexander, 1987), which makes an individual reciprocate the inheritance received from her parents by bequeathing to her children.[1] A second channel is the cultural transmission of preferences (e.g., Bisin and Verdier, 2001). There is a growing literature that endogenizes individuals' preferences by making them a function of their parents' behavior and/or economic conditions (e.g., Rapoport and Vidal, 2007; Adriani and Sonderegger, 2009; Doepke and Zilibotti, 2017). Empirical findings strongly underpin this approach (e.g., Wilhelm *et al.*, 2008; Dohmen *et al.*, 2012). In a recent study, Kosse *et al.* (2020) present evidence on the importance of the social environment (especially the socio-economic status, mother-child interactions and mothers' prosocial attitudes) on the formation of

[1] Arrondel and Masson (2006) classify types of serial indirect reciprocities between family generations. They call this 'backward-downward' indirect reciprocity. Backward refers to the time orientation and means reciprocating an initial act of giving that already took place. Downward refers to the direction of transfers, in this case from parents to children. Kolm (1984) already studied this special type of indirect reciprocity involving at least three generations and called it 'propagation effect'.

prosociality. A third channel is the genetic transmission of preferences. Using a classical twin design, Cesarini *et al.* (2009) estimate that approximately twenty percent of the variation in preferences for giving is explained by genetic differences. Their study relates to the growing behavior genetics literature that finds strong evidence for the heritability of prosociality (e.g., Ebstein *et al.*, 2010; Reuter *et al.*, 2011; Israel *et al.*, 2015; Knafo-Noam *et al.*, 2018; Twito and Knafo-Noam, 2020).

There is some empirical evidence for intergenerational correlation in bequeathing behavior. Arrondel and Grange (2014) use French data to show that the size of bequests left by an individual is more correlated with the size of bequests received than with the remainder of lifetime resources. In earlier studies, Cox and Stark (2005) find similar patterns in U.S. data, as do Arrondel and Masson (2001) as well as Arrondel *et al.* (1997) in French data. Using European data, Stark and Nicinska (2015) find that receiving (or expecting to receive) an inheritance has a positive impact on the intention to bequeath.

This paper discusses the effect of systematic relationships between two generations' preferences for bequeathing on optimal estate tax policy. For this purpose, I set up a non-overlapping two-generation model of parents and children with two types of heterogeneity. First, parents are heterogeneously altruistic. Second, parents inherit differently. One type of parents receives exogenous bequests from the grandparents' generation, the other type does not. The social planner maximizes a (weighted) Utilitarian welfare measure by choosing the optimal estate tax rate and balancing the budget with a lump-sum transfer.

I show that three effects determine the optimal estate tax rate. First, taxing bequests is desirable to redistribute across dynasties depending on welfare weights. Second, positive welfare weights on children increase the optimal tax rate because redistribution also counteracts inequality in this generation. Third, the parental decision to bequeath also yields a positive externality on children that should be corrected by a negative tax (i.e., subsidy). Introducing correlated preferences leads to an additional force towards positive taxation with respect to all three aforementioned effects. Redistributive goals require a higher tax rate and correcting the negative externality is less worthwhile since average bequest sizes are higher. In addition, I find that estate tax rates should be higher for those parents who inherited themselves. In the spirit of Akerlof (1978), this can be thought of as using bequests received as a tag in estate taxation.

This paper relates to the literature on optimal estate and inheritance taxation. Farhi and Werning (2010) and Kopczuk (2013a) consider two-generation models with heterogeneity in parents' productivity, an altruistic or joy-of-giving bequest motive and nonlinear income taxation. In both models, departing from the Atkinson-Stiglitz result of zero estate taxation[2] is optimal if both generations carry a positive welfare weight and non-linear income taxation is available. The optimal tax is progressive since correcting the positive externality from the parental decision to bequeath is costly, and marginal utility from subsidies is decreasing. Kopczuk (2013a) adds an incentive effect to the model. Inheritances decrease the children's labor supply resulting in a decrease of government revenue from income taxation.[3] This fiscal externality is a force for positive estate taxation. Piketty and Saez (2013) consider a dynamic economy with two intra-generational sources of heterogeneity: preferences for bequeathing and productivity. In contrast to previous models, this model hence allows for bi-dimensional inequality of lifetime resources (labor income and inheritances).[4] Nevertheless, in equilibrium, the distributions of bequest sizes and preferences are independent of initial bequest sizes

[2]In these models, one can understand bequests as a consumption good. Then, according to the Atkinson-Stiglitz theorem, taxing bequests is needless if preferences are separable between labor and bequeathing and nonlinear income taxation is available (see Atkinson and Stiglitz, 1976).

[3]See e.g. Kindermann *et al.* (2020), Bø *et al.* (2019), and Elinder *et al.* (2012) for empirical evidence.

[4]A calibration using French and U.S. data suggests the optimal tax rate to be at least 50%.

and preferences due to ergodicity assumptions. My model is closest to the linear tax specification in Farhi and Werning (2013), where parents solely differ in their level of altruism. Compared to their result, bi-dimensional heterogeneity in my model increases the optimal tax rate. None of the preceding models accounts for an intergenerational pattern in bequeathing behavior.

Additionally, this paper contributes to the (mainly income-tax focused) literature on tagging, which conditions tax schedules on other observable characteristics besides earnings. Major contributions use height (Mankiw and Weinzierl, 2010), gender (Alesina *et al.*, 2011; Cremer *et al.*, 2010), sector productivity (Gomes *et al.*, 2018) and age (Bastani *et al.*, 2013; Weinzierl, 2011) as tags. In recent work, Leroux and Pestieau (2020) suggest to condition the taxation of bequests on the age of the deceased. In case of premature death, bequests are at least partly a result of precautionary savings and not completely based on a bequest motive. The optimal differential taxation of early and late bequests in their model depends on both, efficiency and equity considerations. I show that using bequests received by an individual as tags for estate taxation has a positive welfare effect.

The paper proceeds as follows. Section 2 describes the model framework. Section 3 analyzes the optimal estate tax policy for the two cases with and without tagging. Section 4 concludes.

2 Model

I consider a non-overlapping two-generation model where parents and children each live for one period. There are two types of dynasties. In type 1 dynasties, parents receive no bequests from their parents (i.e., the children's grandparents), whereas parents in type 2 dynasties receive a bequest. Parents divide the sum of this bequest (if any) and their exogenous income between consumption and bequeathing to their own children. This decision depends on the parents' degree of altruism, which is heterogenous within both dynasty types and correlates with the dynasty type. Parents have exactly one child that only consumes the bequest received.[5] The government optimizes a (weighted) Utilitarian welfare measure by choosing the estate tax rate and balancing the budget with a lump-sum transfer. Figure (1) illustrates the model set-up.

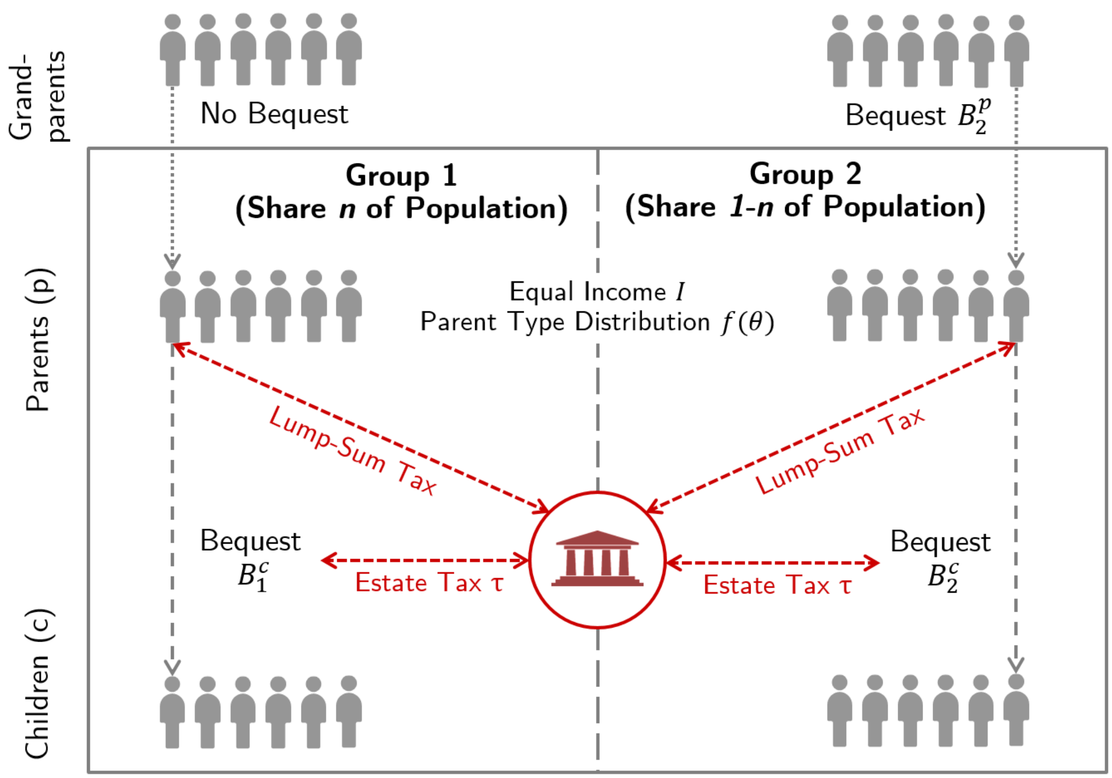

Figure 1: Model Set-Up.

[5]Hence, estate and inheritance taxation coincide. In the case of estate taxation (e.g., in the US), the donor carries the tax liability, which is based on the total estate. In the case of inheritance taxation (e.g., in Germany), the recipient carries the tax liability.

2.1 Individuals

Parents differ in two dimensions: in whether they receive a bequest and in their own taste for bequeathing. Heterogeneity in the first dimension is limited to two different types $g = 1, 2$. All parents of type 1 receive no bequest ($B_1^p = 0$), all parents of type 2 receive a bequest $B_2^p > 0$.[6] The group of type 1 dynasties makes up a share n of the population, the group of type 2 dynasties makes up a share of $1 - n$. Heterogeneity in the second dimension, the degree of altruism, is captured by Θ, which is continuously distributed on the interval $[0, 1]$.

Both dimensions of heterogeneity are systematically linked, i.e., parents who receive an exogenous bequest on average have a higher taste for bequeathing than parents who did not receive a bequest. I assume this for three different reasons. First, receiving a bequest induces reciprocal bequeathing behavior towards the own child. Second, exogenous bequests differ because grandparents are heterogeneously altruistic. The parental generation's preferences, hence, are shaped accordingly through genetic or cultural preference transmission.[7] I model the link between both dimensions by

$$\Theta_g = \theta \cdot a_g(\theta). \tag{1}$$

The preference parameter Θ_g thus consists of two components. The first one, θ, reflects the random part in the taste for bequeathing and is equally distributed within both groups $g = 1, 2$ according to a continuous density function $f(\theta)$ defined on the interval

[6] Results remain qualitatively the same if type 1 parents receive a positive bequest B_1^p that is smaller than B_2^p.

[7] This, of course, means that I abstract from productivity and wealth differences in the grandparents' generation.

$[\theta_{low}, \theta_{high}]$. The second one, $a_g(\theta)$, links both dimensions of heterogeneity with

$$0 < a_1(\theta) < 1 \quad \text{and} \quad a_2(\theta) > 1 \quad \forall\theta. \tag{2}$$

Due to these characteristics of the second preference term, type 2 parents are on average more altruistic than type 1 parents. Since a is a function of θ, not only the average level of altruism may be different in both groups, but also the shape of the distribution of Θ. This captures the intergenerational correlation of preferences for bequeathing within different dynasties.

Parents receive utility u_g^p from consumption and from leaving a bequest to their child. For type g, parents' utility is given by

$$u_g^p(C_g^p, C_g^c; \Theta_g) = (1 - \Theta_g)\ln(C_g^p) + \Theta_g \ln(C_g^c), \tag{3}$$

where C_g^p is the parents' consumption and C_g^c is the child's consumption. By using this utility function, I assume additive separability between consumption and bequests. All parents receive equal exogenous income I, type 2 parents also receive exogenous bequests B_2^p from their own parents. Given a linear savings technology with a periodical gross rate of return R and a linear estate tax rate τ, the budget constraint for parents of type g is then

$$C_g^p + \frac{C_g^c}{R(1-\tau)} = I + B_g^p, \tag{4}$$

where $\frac{C_g^c}{R(1-\tau)}$ amounts to the gross bequest left. In the following, I use $\hat{R} = R(1-\tau)$ for ease of notation. Solving the parents' optimization problem, i.e., maximizing (3) subject to the budget constraint (4), yields the indirect utility function

$$v_g^p(I, B_g^p, \hat{R}; \Theta_g) = (1 - \Theta_g) \ln \left[(I + B_g^p)(1 - \Theta_g)\right] + \Theta_g \ln \left[(I + B_g^p)\Theta_g \hat{R}\right]. \quad (5)$$

Due to cross price elasticities of zero, the estate tax only enters in the last term, i.e., it only changes the net-of-tax bequest. Consumption and gross bequests are perfectly inelastic with respect to the tax rate. Equation (5) also implies that, before redistribution, parents receiving a bequest have lower marginal utility from their funds compared to parents not receiving a bequest.

Children only consume and have utility

$$u_g^c(C_g^c) = \ln(C_g^c) = \ln\left((I + B_g^p)\Theta_g \hat{R}\right). \quad (6)$$

Children of type 1 dynasties and children of type 2 dynasties differ in their utility even when they have equally altruistic parents. Parents in both groups bequeath the same share of their funds, however, the absolute fund size differs due to the differences in bequests received from the grandparents.

2.2 Government

The government maximizes a (weighted) Utilitarian welfare measure given by

$$W = \int_{\theta_{low}}^{\theta_{high}} \left[n\lambda_1 \left(v_1^p + \delta u_1^c\right) + (1-n)\lambda_2 \left(v_2^p + \delta u_2^c\right)\right] f(\theta) d\theta, \quad (7)$$

where λ_1, λ_2 are welfare weights on type 1 and type 2 dynasties and $0 \leq \delta \leq 1$ is a generational discount factor.[8] The first term represents aggregated welfare of all type 1 dynasties, the second term represents aggregated welfare of all type 2 dynasties. The resource constraint

$$E = \int_{\theta_{low}}^{\theta_{high}} \left[n \left(C_1^p + \frac{1}{r} C_1^c \right) + (1-n) \left(C_2^p + \frac{1}{r} C_2^c \right) \right] f(\theta) d\theta \leq \bar{E} \tag{8}$$

with $\bar{E}$ being the maximum level of resources available restricts the government. This expression corresponds to the total exogenous income in the economy.

The government does not observe the preferences of the individuals, but solely their consumption behavior and bequest sizes B_g^p, B_g^c. Hence, the first-best solution is unavailable. I restrict policy instruments to a linear estate tax (or subsidy) and a lump-sum tax (or transfer). The government solves

$$\max_{\hat{R}, I} \mathcal{L} = W(I, B_2^p, \hat{R}; \Theta_g) - \kappa \left[E - \bar{E} \right], \tag{9}$$

where κ denotes the Lagrangian multiplier. Variable $\hat{R} = R(1-\tau)$ captures the estate tax rate, I captures the lump-sum tax (or transfer). Since income is equal for all individuals in the parental generation, the lump-sum tax (or transfer) is equivalent to a linear income tax (or transfer).

[8] A positive discount factor means that the government explicitly considers the children's welfare. This may seem problematic from a political economy perspective since future generations do not elect the government. However, the parents' voting decisions may already (partly) reflect their children's perspective. For some general discussion on social discounting see Farhi and Werning (2007). In the context of voluntary bequests, including both generations in the welfare function leads to double-counting the benefits of bequests (since both donor and donee receive utility). While this approach has become standard in the literature, Boadway and Cuff (2015) critically discuss its implications.

3 Optimal Policy

3.1 Linear Estate Tax without Tagging

In a first step, I consider a unique tax rate that applies to all parents, i.e., the government does not condition estate taxes on bequests received. This specification is close to what one observes in reality to date. Solving the maximization problem described in section 2.2 gives the optimal estate tax rate and lump-sum tax. In the following, I omit the lower bound (θ_{low}) and upper bound (θ_{high}) of integrals to simplify notation. Also, I denote the budget-weighted population average of the preference parameter as

$$0 \ < \ \bar{\Theta}_w = nI \int \Theta_1 f(\theta) d\theta + (1-n)(I + B_2^p) \int \Theta_2 f(\theta) d\theta. \tag{10}$$

This corresponds to average bequests before taxes and interest. The unweighted population average is

$$0 \ < \ \bar{\Theta} = n \int \Theta_1 f(\theta) d\theta + (1-n) \int \Theta_2 f(\theta) d\theta \ < \ 1. \tag{11}$$

PROPOSITION 1: *When $\kappa(\lambda_1, \lambda_2, \delta)$ denotes the Lagrangian multiplier on the governmental resource constraint, the optimal linear estate tax rate in the case of equal welfare weights ($\lambda_1 = \lambda_2 = 1$) is*

$$\tau = \frac{(1+\delta)(n^2 - n)}{\kappa} \frac{\int \Theta_1 f(\theta) d\theta \left(1 - \frac{I}{I + B_2^p}\right) + \int \Theta_2 f(\theta) d\theta \left(1 - \frac{I + B_2^p}{I}\right)}{(1 - \bar{\Theta}) \bar{\Theta}_w} - \frac{\delta}{\kappa} \frac{1}{\bar{\Theta}_w}.$$

If welfare weights differ between groups, the optimal tax rate is

$$\begin{aligned}\tau =&\frac{(1+\delta)(n^2-n)}{\kappa}\frac{\int \Theta_1 f(\theta)d\theta\left(\lambda_1-\lambda_2\frac{I}{I+B_2^p}\right)+\int \Theta_2 f(\theta)d\theta\left(\lambda_2-\lambda_1\frac{I+B_2^p}{I}\right)}{(1-\bar{\Theta})\bar{\Theta}_w}\\ &-\frac{\delta}{\kappa}\frac{n\lambda_1\left(1-\int \Theta_1 f(\theta)d\theta\right)+(1-n)\lambda_2\left(1-\int \Theta_2 f(\theta)d\theta\right)}{(1-\bar{\Theta})\bar{\Theta}_w}.\end{aligned}$$

PROOF: see Appendix A.1.

Three effects determine the result in the first part of Proposition 1. First, assume that the welfare function does not include the children's generation, i.e., $\delta = 0$. Then, the second term drops out and the optimal estate tax is unambiguously positive.[9] The equation hence calls for taxing bequests and as a consequence for lump-sum transfers. This result reflects the force for redistribution across dynasties. A positive estate tax raises more revenue from type 2 dynasties than from type 1 dynasties. The government redistributes this revenue through a lump-sum transfer to all dynasties. As a consequence, overall welfare increases since disposable income (with decreasing marginal utility) becomes more equal across different dynasty types.

Second, if there is positive weight δ on the children's generation, the first term in the optimal tax formula increases by factor δ. Equalizing parents' disposable income also leads to more similar bequest sizes. Since children's marginal utility of consumption is decreasing, this improves welfare in the children's generation.

Third, the second term in the formula represents a Pigouvian subsidy on bequests. The social welfare function reflects the parental utility from giving through v_g^p and the benefit for the children (namely the increase in budget) through u_g^c. Hence, the

[9] To see this, note that the denominator of the remaining term is positive because of (10) and (11). The numerator is also unambiguously positive for $0 < n < 1$. This is because $\int \Theta_2 f(\theta)d\theta > \int \Theta_1 f(\theta)d\theta$ and $\left(1-\frac{I}{I+B_2^p}\right) < \left|1-\frac{I+B_2^p}{I}\right|$, which is equivalent to $\frac{1}{I+B_2^p} < \frac{1}{I}$.

optimal bequeathing decision from the parents' individual perspective is not socially optimal; bequeathing yields a positive externality on children. This externality calls for a corrective subsidy. The overall sign of the optimal tax rate depends on the relative size of the two terms, since the first one is positive and the second one is negative.

In the second part of Proposition 1, I allow for different welfare weights for different dynasty types. If welfare weights are such that the social marginal utility from income is higher for type 2 parents than for type 1 parents, the first term would turn negative, calling for subsidies on bequests not only from a Pigouvian but also from a redistributive perspective. However, higher welfare weights on individuals receiving bequests seem hard to justify.

Now, how does the correlation in preferences for bequeathing influence the optimal tax rate?

PROPOSITION 2: *Let* τ_{pos} (> 0) *denote the first term and* τ_{neg} (< 0) *denote the second term in the optimal tax formula. Assume a parallel increase in* $\int \Theta_2 f(\theta)d\theta$ *and decrease in* $\int \Theta_1 f(\theta)d\theta$ *such that* $\bar{\Theta} = c = const.$ *with* $0 < c < 1$*. For equal welfare weights* $\lambda_1 = \lambda_2 = 1$ *it holds that*

$$\left.\frac{\partial \tau}{\partial(\int \Theta_2 f(\theta)d\theta)}\right|_{\bar{\Theta}=c} > 0, \quad \left.\frac{\partial \tau_{pos}}{\partial(\int \Theta_2 f(\theta)d\theta)}\right|_{\bar{\Theta}=c} > 0, \quad \left.\frac{\partial \tau_{neg}}{\partial(\int \Theta_2 f(\theta)d\theta)}\right|_{\bar{\Theta}=c} > 0. \quad (12)$$

A stronger correlation in bequest motives implies a higher optimal tax rate. The positive redistributive term increases in magnitude and the negative Pigouvian term decreases in absolute terms.

PROOF: see Appendix A.2.

This parallel change in the distributions of the bequest motive within the two groups implies a mean-preserving spread in the distribution of the bequest motive across the whole population. While the population average of the bequest motive remains constant, the motive becomes more correlated across dynasties. The overall tax rate increases, but still can be positive or negative. Remarkably, both terms in the optimal tax formula increase as the two other derivatives in Proposition 2 show.

The positive first term reflects the force for redistribution, which increases welfare of type 1 dynasties. However, this positive tax part also hurts type 1 parents by increasing the price for bequeathing. A decrease in the preference for bequeathing in these dynasties reduces the cost of redistribution, i.e., reduces the utility loss due to a higher price. This cost reduction overcompensates the additional cost of redistribution in type 2 dynasties that arise due to higher preferences for bequeathing. This overcompensation is due to the (on average) higher bequests in type 2 dynasties and the corresponding decrease in marginal utilities from bequeathing.

The increase in the negative second term, which means a decrease in absolute terms, is more straightforward. The average net-of-tax bequest size

$$\bar{B}^c = nI\hat{R}\int \Theta_1 f(\theta)d\theta + (1-n)(I + B_2^p)\hat{R}\int \Theta_2 f(\theta)d\theta$$

increases due to the parallel change in preferences, i.e.

$$\left.\frac{\partial \bar{B}_c}{\partial(\int \Theta_2 f(\theta)d\theta)}\right|_{\bar{\Theta}=c;\ \hat{R}=const.} = \hat{R}(1-n)B_2^p \ > \ 0.$$

Hence, subsidizing bequests becomes less worthwhile on average when the correlation in bequeathing behavior across generations is more pronounced. This result for the Pigouvian term is easily transferrable to different optimal estate tax models, where the estate tax is not the sole instrument for redistribution.

3.2 Linear Estate Tax with Tagging

If the government observes bequests received and bequests left, the tax system could discriminate between type 1 and type 2 dynasties. The linear estate tax then applies different tax rates to the respective dynasties while the government still balances its budget with a unique lump-sum tax or transfer.[10]

PROPOSITION 3: *Let τ_1 denote the optimal tax rate for type 1 dynasties and τ_2 the optimal tax rate for type 2 dynasties. It holds that welfare-maximizing tax rates under the weighted utilitarian welfare measure fulfil*

$$\tau_2 > \tau_1 \quad \text{iff} \quad \frac{\lambda_1}{\lambda_2} > \frac{\frac{1}{I+B_2^p} + \frac{\delta}{(I+B_2^p)\int \Theta_2 f(\theta)d\theta}}{\frac{1}{I} + \frac{\delta}{I\int \Theta_1 f(\theta)d\theta}} = \zeta.$$

PROOF: see Appendix A.3.

When allowing for different tax rates, the sign of the tax rate differential depends on the (pre-tax) average social marginal utility of income in the two groups of dynasty types. The numerator of the condition in Proposition 3 shows the sum of the parents' marginal utility of income and the children's (discounted) average marginal utility of income for type 2 dynasties. The denominator shows the same for type 1 dynasties.

Hence, bequests in type 2 dynasties should be taxed at a higher rate if their average social marginal utility of income is smaller than in type 1 dynasties. This is definitely true for equal welfare weights and for $\lambda_1 > \lambda_2$ as type 2 parents have more funds due to bequests received and type 2 children receive higher bequests on average. Additionally, this even holds if more weight is put on dynasties where parents receive a bequest as

[10] In a more general setting with some continuous distribution over dynasty types, the applicable estate tax rate would be a function of bequests received by the same individual.

long as $\lambda_1/\lambda_2 > \zeta$ holds.

3.3 Discussion

The model assumes that gross bequests are inelastic with respect to the estate tax rate. While this makes the model more tractable, it may neglect a potential response to estate taxation. In a model where parents solely differ in their level of altruism, Farhi and Werning (2013) allow for elastic gross bequests through a more general utility function. Concerning the redistribution across dynasties, this adds the inverse-elasticity rule to the optimal tax formula, which reflects the cost and benefits of intragenerational redistribution. If gross bequests are more elastic, the optimal tax rate decreases. Concerning the redistribution across generations, i.e. the Pigouvian tax motive, elasticities only work as weights for the children's social marginal utility. Applied to my model, the impact on the optimal taxation of bequests would not only be a question of how strongly gross bequests respond to taxation on average, but also of how elasticities differ across the two groups.

By considering the initial bequests (i.e. the ones from grandparents to parents) as exogenous, the dynamic character of the model is somewhat limited. These bequests cannot be taxed in the model to already redistribute at that stage. One can think about that as considering only the short run effect over the next generation where bequests have already been received and cannot be taxed anymore. However, even over the period of several generations the key mechanism would persist. As long as some heterogeneity in after-tax bequests remains, path-dependency of bequest motives would lead to an accumulation of inequality with respect to bequests received over time. The degree to which taxation can limit this accumulation depends on the relative importance of the three correlation channels. Whereas preference transmission (genetic and cultural) is independent of whether the bequest motive is operative (and hence also independent of the tax on bequests), indirect reciprocal behavior depends on actually receiving a bequest. If reciprocity depends on the net-of-tax bequest, taxation could indeed decrease the effect of this channel. However, reciprocity may be driven by the intention of the bequeather and therefore depend on gross bequests. This is why a richer model could look at the three channels separately, whereas they are all captured by one parameter in my model.

4 Conclusion

In the last decade, the theoretical literature on optimal estate or inheritance taxation has discussed the main drivers of optimal tax rates. In this paper, I have added another determinant of optimal taxation to the discussion: intergenerational correlation in bequeathing behavior, which arises due to genetic or cultural transmission of preferences for bequeathing or due to indirect reciprocity. The optimal tax formula trades off a redistributive force for positive taxation and a force for negative taxation to correct the positive externality from giving. This general result is as in Farhi and Werning (2013). In addition, first, I show that a stronger relationship between bequests received and the willingness to leave bequests results in an increase in the optimal tax rate. Second, I show that when allowing for different tax rates across dynasties, the ones with higher bequests received should face a higher tax rate. Hence, bequests received should serve as "tags" in estate taxation when abstracting from additional administrative cost.

From the perspective of policy makers, my second contribution is of immediate interest. My results indicate that it is welfare-increasing if estate tax rates are based on the size of bequests received by the same individual. In this sense, it matters whether inequality in lifetime resources arises due to heterogeneity in productivity or due to heterogeneity in bequests received.

References

Adermon, A., Lindahl, M. and Waldenström, D. (2018). Intergenerational Wealth Mobility and the Role of Inheritance: Evidence from Multiple Generations. *Economic Journal*, **128** (612), F482–F513.

Adriani, F. and Sonderegger, S. (2009). Why Do Parents Socialize Their Children to Behave Pro-Socially? An Information-Based Theory. *Journal of Public Economics*, **93** (11-12), 1119–1124.

Akerlof, G. A. (1978). The Economics of "Tagging" as Applied to the Optimal Income Tax, Welfare Programs, and Manpower Planning. *American Economic Review*, **68** (1), 8–19.

Alesina, A., Ichino, A. and Karabarbounis, L. (2011). Gender-Based Taxation and the Division of Family Chores. *American Economic Journal: Economic Policy*, **3** (2), 1–40.

Alexander, R. D. (1987). *The Biology of Moral Systems*. New York: De Gruyter.

Arrondel, L. and Grange, C. (2014). Bequests and Family Traditions: The Case of Nineteenth Century France. *Review of Economics of the Household*, **12** (3), 439–459.

— and Masson, A. (2001). Family Transfers Involving Three Generations. *The Scandinavian Journal of Economics*, **103** (1033), 415–443.

— and — (2006). Altruism, Exchange or Indirect Reciprocity: What Do the Data on Family Transfers Show? In S.-C. Kolm and J. M. Ythier (eds.), *Handbook of the Economics of Giving, Altruism and Reciprocity*, vol. 2, *14*, Amsterdam et al.: Elsevier, pp. 971–1053.

—, — and PESTIEAU, P. (1997). Bequest and Inheritance: Empirical Issues and France-U.S. Comparison. In G. Erreygers and T. Vandevelde (eds.), *Is Inheritance Legitimate?*, *4*, Berlin et al.: Springer, pp. 89–125.

ATKINSON, A. B. and STIGLITZ, J. E. (1976). The Design of Tax Structure: Direct versus Indirect Taxation. *Journal of Public Economics*, **6** (1-2), 55–75.

BASTANI, S., BLOMQUIST, S. and MICHELETTO, L. (2013). The Welfare Gains of Age-Related Optimal Income Taxation. *International Economic Review*, **54** (4), 1219–1249.

BISIN, A. and VERDIER, T. (2001). The Economics of Cultural Transmission and the Dynamics of Preferences. *Journal of Economic Theory*, **97** (2), 298–319.

BØ, E. E., HALVORSEN, E. and THORESEN, T. O. (2019). Heterogeneity of the Carnegie Effect. *Journal of Human Resources*, **54** (3), 726–759.

BOADWAY, R. and CUFF, K. (2015). Tax Treatment of Bequests when Donor Benefits Are Discounted. *International Tax and Public Finance*, **22** (4), 604–634.

CESARINI, D., DAWES, C. T., JOHANNESSON, M., LICHTENSTEIN, P. and WALLACE, B. (2009). Genetic Variation in Preferences for Giving and Risk Taking. *Quarterly Journal of Economics*, **124** (2), 809–842.

COX, D. and STARK, O. (2005). Bequests, Inheritances and Family Traditions. *CRR Working Paper*, **2005-09**.

CREMER, H., GAHVARI, F. and LOZACHMEUR, J.-M. (2010). Tagging and Income Taxation: Theory and an Application. *American Economic Journal: Economic Policy*, **2** (1), 31–50.

Doepke, M. and Zilibotti, F. (2017). Parenting With Style: Altruism and Paternalism in Intergenerational Preference Transmission. *Econometrica*, **85** (5), 1331–1371.

Dohmen, T., Falk, A., Huffman, D. and Sunde, U. (2012). The Intergenerational Transmission of Risk and Trust Attitudes. *Review of Economic Studies*, **79** (2), 645–677.

Ebstein, R. P., Israel, S., Chew, S. H., Zhong, S. and Knafo, A. (2010). Genetics of Human Social Behavior. *Neuron*, **65** (6), 831–844.

Elinder, M., Erixson, O. and Ohlsson, H. (2012). The Impact of Inheritances on Heirs' Labor and Capital Income. *B.E. Journal of Economic Analysis and Policy: Contributions*, **12** (1), 1–37.

Farhi, E. and Werning, I. (2007). Inequality and Social Discounting. *Journal of Political Economy*, **115** (3), 365–402.

— and — (2010). Progressive Estate Taxation. *The Quarterly Journal of Economics*, **125** (2), 635–673.

— and — (2013). Estate Taxation with Altruism Heterogeneity. *American Economic Review: Papers & Proceedings*, **103** (3), 489–495.

Gomes, R., Lozachmeur, J. M. and Pavan, A. (2018). Differential Taxation and Occupational Choice. *Review of Economic Studies*, **85** (1), 511–557.

Israel, S., Hasenfratz, L. and Knafo-Noam, A. (2015). The Genetics of Morality and Prosociality. *Current Opinion in Psychology*, **6**, 55–59.

Kindermann, F., Mayr, L. and Sachs, D. (2020). Inheritance Taxation and Wealth

Effects on the Labor Supply of Heirs. *Journal of Public Economics*, **191**, Article No. 104127.

Knafo-Noam, A., Vertsberger, D. and Israel, S. (2018). Genetic and Environmental Contributions to Children's Prosocial Behavior: Brief Review and New Evidence from a Reanalysis of Experimental Twin Data. *Current Opinion in Psychology*, **20**, 60–65.

Kolm, S.-C. (1984). La Bonne Economie: la Réciprocité Générale. *Presses universitaires de France*, **33**.

Kopczuk, W. (2013a). Incentive Effects of Inheritances and Optimal Estate Taxation. *American Economic Review: Papers & Proceedings*, **103** (3), 472–477.

— (2013b). Taxation of Intergenerational Transfers and Wealth. In A. Auerbach, R. Chetty, M. Feldstein and E. Saez (eds.), *Handbook of Public Economics*, vol. 5, *6*, Amsterdam et al.: Elsevier B.V., pp. 329–390.

Kosse, F., Deckers, T., Pinger, P., Schildberg-Hörisch, H. and Falk, A. (2020). The Formation of Prosociality: Causal Evidence on the Role of Social Environment. *Journal of Political Economy*, **128** (2), 434–467.

Leroux, M.-L. and Pestieau, P. (2020). Age-Related Taxation of Bequests in the Presence of a Dependency Risk. *CESifo Working Paper Series*, **8642**.

Mankiw, N. G. and Weinzierl, M. (2010). The Optimal Taxation of Height: A Case Study of Utilitarian Income Redistribution. *American Economic Journal: Economic Policy*, **2** (1), 155–176.

Piketty, T. (2011). On the Long-Run Evolution of Inheritance: France 1820-2050. *Quarterly Journal of Economics*, **126** (3), 1071–1131.

— and Saez, E. (2013). A Theory of Optimal Inheritance Taxation. *Econometrica*, **81** (5), 1851–1886.

Rapoport, H. and Vidal, J. P. (2007). Economic Growth and Endogenous Intergenerational Altruism. *Journal of Public Economics*, **91** (7-8), 1231–1246.

Reuter, M., Frenzel, C., Walter, N. T., Markett, S. and Montag, C. (2011). Investigating the Genetic Basis of Altruism: The Role of the COMT Val158Met Polymorphism. *Social Cognitive and Affective Neuroscience*, **6** (5), 662–668.

Stark, O. and Nicinska, A. (2015). How Inheriting Affects Bequest Plans. *Economica*, **82** (s1), 1126–1152.

Twito, L. and Knafo-Noam, A. (2020). Beyond culture and the family: Evidence from twin studies on the genetic and environmental contribution to values. *Neuroscience and Biobehavioral Reviews*, **112**, 135–143.

Weinzierl, M. (2011). The Surprising Power of Age-Dependent Taxes. *Review of Economic Studies*, **78** (4), 1490–1518.

Wilhelm, M. O., Brown, E., Rooney, P. M. and Steinberg, R. (2008). The Intergenerational Transmission of Generosity. *Journal of Public Economics*, **92** (10-11), 2146–2156.

A Appendix

A.1 Proof of Proposition 1

To solve the governmental planning problem, I study the Lagrangian

$$\mathcal{L} = W(I, B_2^p, \hat{R}; \Theta_g) - \kappa \left[E - \bar{E}\right] \tag{A.1}$$

with

$$\begin{aligned} W &= \int_{\theta_{low}}^{\theta_{high}} \left[n\lambda_1 \left(v_1^p + \delta u_1^c\right) + (1-n)\lambda_2 \left(v_2^p + \delta u_2^c\right)\right] f(\theta) d\theta, \\ E &= \int_{\theta_{low}}^{\theta_{high}} \left[n \left(C_1^p + \frac{1}{R} C_1^c\right) + (1-n) \left(C_2^p + \frac{1}{R} C_2^c\right)\right] f(\theta) d\theta. \end{aligned}$$

The first-order conditions are

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \hat{R}} =& \frac{1}{\hat{R}} \left[n\lambda_1 \left(\int \Theta_1 f(\theta) d\theta + \delta\right) + (1-n)\lambda_2 \left(\int \Theta_2 f(\theta) d\theta + \delta\right)\right] \\ & - \kappa \frac{1}{r} \left[nI \int \Theta_1 f(\theta) d\theta + (1-n)(I + B_2^p) \int \Theta_2 f(\theta) d\theta\right] = 0, \end{aligned} \tag{A.2}$$

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial I} =& \frac{n\lambda_1 (1+\delta)}{I} + \frac{(1-n)\lambda_2 (1+\delta)}{I + B_2^p} \\ & - \kappa \left[n \left(\int (1-\Theta_1) f(\theta) d\theta + \frac{\hat{R}}{R} \int \Theta_1 f(\theta) d\theta\right)\right. \\ & \left. + (1-n) \left(\int (1-\Theta_2) f(\theta) d\theta + \frac{\hat{R}}{R} \int \Theta_2 f(\theta) d\theta\right)\right] = 0. \end{aligned} \tag{A.3}$$

Equations (A.2) and (A.3) can be rearranged to

$$\frac{\hat{R}}{R} = \frac{n\lambda_1 \int \Theta_1 f(\theta) d\theta + (1-n)\lambda_2 \int \Theta_2 f(\theta) d\theta + \delta \left(n\lambda_1 + (1-n)\lambda_2\right)}{\kappa \bar{\Theta}_w}, \tag{A.4}$$

$$\frac{\hat{R}}{R} = \frac{\frac{n\lambda_1(1+\delta)}{I} + \frac{(1-n)\lambda_2(1+\delta)}{I+B_2^p} - \kappa\left(1-\bar{\Theta}\right)}{\kappa\bar{\Theta}}, \tag{A.5}$$

respectively. Combining equations (A.4) and (A.5) and solving for κ yields

$$\begin{aligned}\kappa = &- \frac{\bar{\Theta}\left[n\lambda_1 \int \Theta_1 f(\theta)d\theta + (1-n)\lambda_2 \int \Theta_2 f(\theta)d\theta + \delta\left(n\lambda_1 + (1-n)\lambda_2\right)\right]}{\bar{\Theta}_w\left(1-\bar{\Theta}\right)} \\ &+ \frac{(1+\delta)\left[\frac{n\lambda_1}{I} + \frac{(1-n)\lambda_2}{I+B_2^p}\right]}{1-\bar{\Theta}}.\end{aligned} \tag{A.6}$$

After reinserting expression (A.6) into the numerator of equation (A.5) and some further manipulations, I obtain

$$\tau = -\frac{1}{\kappa}\frac{n\lambda_1 \int \Theta_1 f(\theta)d\theta + (1-n)\lambda_2 \int \Theta_2 f(\theta)d\theta + \delta\left(n\lambda_1 + (1-n)\lambda_2\right) - \kappa\bar{\Theta}_w}{\bar{\Theta}_w}. \tag{A.7}$$

Inserting expression (A.6) into the numerator in (A.7) and simplifying gives the equation in Proposition 1.

■

A.2 Proof of Proposition 2

In Proposition 2, I define

$$\tau_{pos} = \frac{(1+\delta)(n^2-n)}{\kappa}\frac{\int \Theta_1 f(\theta)d\theta\left(1-\frac{I}{I+B_2^p}\right) + \int \Theta_2 f(\theta)d\theta\left(1-\frac{I+B_2^p}{I}\right)}{(1-\bar{\Theta})\bar{\Theta}_w} \tag{A.8}$$

and

$$\tau_{neg} = -\frac{\delta}{\kappa}\frac{1}{\bar{\Theta}_w}. \tag{A.9}$$

A parallel increase in $\int \Theta_2 f(\theta)d\theta$ and decrease in $\int \Theta_1 f(\theta)d\theta$ with $\bar{\Theta} = c = const.$ implies

$$\int \Theta_1 f(\theta)d\theta = \frac{c - (1-n)\int \Theta_2 f(\theta)d\theta}{n}. \tag{A.10}$$

Inserting equations (A.6) and (A.10) into (A.8) and (A.9) gives

$$\bar{\tau}_{pos} = \frac{(1+\delta)(n^2-n)\left[\frac{c-(1-n)\int\Theta_2 f(\theta)d\theta}{n}\left(1-\frac{I}{I+B_2^p}\right)+\int\Theta_2 f(\theta)d\theta\left(1-\frac{I+B_2^p}{I}\right)\right]}{\left(\int\Theta_2 f(\theta)d\theta(1-n)B_2^p+cI\right)\left(\frac{n(1+\delta)}{I}+\frac{(1-n)(1+\delta)}{I+B_2^p}\right)-c\,(c+\delta)} \tag{A.11}$$

and

$$\bar{\tau}_{neg} = -\frac{\delta(1-c)}{\left(\int\Theta_2 f(\theta)d\theta(1-n)B_2^p+cI\right)\left(\frac{n(1+\delta)}{I}+\frac{(1-n)(1+\delta)}{I+B_2^p}\right)-c\,(c+\delta)}. \tag{A.12}$$

The first derivatives of (A.11) and (A.12) are

$$\frac{\partial\bar{\tau}_{pos}}{\partial(\int\Theta_2 f(\theta)d\theta)} = \left.\frac{\partial\tau_{pos}}{\partial(\int\Theta_2 f(\theta)d\theta)}\right|_{\bar{\Theta}=c} =$$

$$\frac{(1+\delta)(n^2-n)\left[-\frac{1-n}{n}\left(1-\frac{I}{I+B_p^2}\right)+\left(1-\frac{I+B_p^2}{I}\right)\right]c(1-c)}{\left(\left(\int\Theta_2 f(\theta)d\theta(1-n)B_2^p+cI\right)\left(\frac{n(1+\delta)}{I}+\frac{(1-n)(1+\delta)}{I+B_2^p}\right)-c\,(c+\delta)\right)^2} \tag{A.13}$$

and

$$\frac{\partial\bar{\tau}_{neg}}{\partial(\int\Theta_2 f(\theta)d\theta)} = \left.\frac{\partial\tau_{neg}}{\partial(\int\Theta_2 f(\theta)d\theta)}\right|_{\bar{\Theta}=c} =$$

$$\frac{\delta(1+\delta)(n^2-n)\left[-\frac{1-n}{n}\left(1-\frac{I}{I+B_p^2}\right)+\left(1-\frac{I+B_p^2}{I}\right)\right](1-c)}{\left(\left(\int\Theta_2 f(\theta)d\theta(1-n)B_2^p+cI\right)\left(\frac{n(1+\delta)}{I}+\frac{(1-n)(1+\delta)}{I+B_2^p}\right)-c\,(c+\delta)\right)^2}. \tag{A.14}$$

The denominators of both derivatives are positive. Since $c, n < 1$ and $B_p^2 > 0$, the numerators are also positive. From these two derivatives being positive, it follows immediately that

$$\left.\frac{\partial \tau}{\partial(\int \Theta_2 f(\theta)d\theta)}\right|_{\bar{\theta}=c} > 0. \tag{A.15}$$

■

A.3 Proof of Proposition 3

With different estate tax rates for different dynasty types, the indirect utility functions transform to

$$v_g^p(I, B_g^p, \hat{R}_g; \Theta_g) = (1-\Theta_g)\ln\left[(I+B_g^p)(1-\Theta_g)\right] + \Theta_g \ln\left[(I+B_g^p)\Theta_g \hat{R}_g\right].$$

I solve the adapted Lagrangian $\mathcal{L} = \tilde{W}(I, B_2^p, \hat{R}_g; \Theta_g) - \kappa\left[E - \bar{E}\right]$.

The first-order conditions are

$$\frac{\partial \mathcal{L}}{\partial \hat{R}_1} = 0 \quad \Leftrightarrow \quad \frac{\hat{R}_1}{R} = \frac{\lambda_1\left(\int \Theta_1 f(\theta)d\theta + \delta\right)}{\kappa I \int \Theta_1 f(\theta)d\theta}, \tag{A.16}$$

$$\frac{\partial \mathcal{L}}{\partial \hat{R}_2} = 0 \quad \Leftrightarrow \quad \frac{\hat{R}_2}{R} = \frac{\lambda_2\left(\int \Theta_2 f(\theta)d\theta + \delta\right)}{\kappa (I+B_2^p) \int \Theta_2 f(\theta)d\theta}, \tag{A.17}$$

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial I} =& \frac{n\lambda_1(1+\delta)}{I} + \frac{(1-n)\lambda_2(1+\delta)}{I+B_2^p} \\ &- \kappa\left[n\left(1+\left(\frac{\hat{R}_1}{R}-1\right)\int \Theta_1 f(\theta)d\theta\right)\right. \\ &\left.+(1-n)\left(1+\left(\frac{\hat{R}_2}{R}-1\right)\int \Theta_2 f(\theta)d\theta\right)\right] = 0\end{aligned} \tag{A.18}$$

Solving equations (A.16) and (A.17) for the optimal tax rate gives

$$\tau_1 = \frac{\kappa I \int \Theta_1 f(\theta) d\theta - \lambda_1 (\int \Theta_1 f(\theta) d\theta + \delta)}{\kappa I \int \Theta_1 f(\theta) d\theta}, \tag{A.19}$$

$$\tau_2 = \frac{\kappa (I + B_2^p) \int \Theta_2 f(\theta) d\theta - \lambda_2 (\int \Theta_2 f(\theta) d\theta + \delta)}{\kappa (I + B_2^p) \int \Theta_2 f(\theta) d\theta}. \tag{A.20}$$

Inserting equations (A.16) and (A.17) into (A.18) yields an expression for κ. After inserting this expression into the numerators of equations (A.19) and (A.20), and after some further manipulations, I get the optimal tax rate formulas

$$\tau_1 = \frac{1-n}{\kappa} \frac{\left(\lambda_2 \frac{1}{I+B_2^p} - \lambda_1 \frac{1}{I}\right)(1 - \int \Theta_2 f(\theta) d\theta)}{1 - \bar{\Theta}} - \frac{\lambda_1 \delta}{\kappa} \frac{1}{I \int \Theta_1 f(\theta) d\theta}, \tag{A.21}$$

$$\tau_2 = \frac{n}{\kappa} \frac{\left(\lambda_1 \frac{1}{I} - \lambda_2 \frac{1}{I+B_2^p}\right)(1 - \int \Theta_1 f(\theta) d\theta)}{1 - \bar{\Theta}} - \frac{\lambda_2 \delta}{\kappa} \frac{1}{(I + B_2^p) \int \Theta_2 f(\theta) d\theta}. \tag{A.22}$$

Calculating the difference in optimal tax rates between different dynasty types yields

$$\tau_2 - \tau_1 = \frac{\lambda_1 \frac{1}{I} - \lambda_2 \frac{1}{I+B_2^p}}{\kappa} + \frac{\delta}{\kappa} \left(\frac{\lambda_1}{I \int \Theta_1 f(\theta) d\theta} - \frac{\lambda_2}{(I + B_2^p) \int \Theta_2 f(\theta) d\theta} \right). \tag{A.23}$$

Rearranging (A.23) yields the equation in Proposition 3. Since $B_2^p > 0$ and $\int \Theta_2 f(\theta) d\theta > \int \Theta_1 f(\theta) d\theta$, for $\lambda_1 = \lambda_2$ it holds that $\tau_2 > \tau_1$.

■