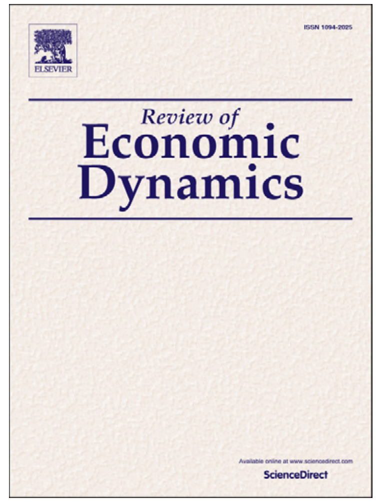
Accepted Manuscript



Transitional dynamics and the optimal progressivity of income redistribution

Ozan Bakış, Barış Kaymak, Markus Poschke

PII: S1094-2025(14)00048-9
DOI: 10.1016/j.red.2014.08.004
Reference: YREDY 690

To appear in: *Review of Economic Dynamics*

Received date: 22 October 2013
Revised date: 22 July 2014

Please cite this article in press as: Bakış, O., et al. Transitional dynamics and the optimal progressivity of income redistribution. *Review of Economic Dynamics* (2014), http://dx.doi.org/10.1016/j.red.2014.08.004

Transitional Dynamics and the Optimal Progressivity of Income Redistribution

Ozan Bakış[a], Barış Kaymak[b], Markus Poschke[c]

[a]TUSIAD-Sabanci University Competitiveness Forum, Istanbul, Turkey
[b]Department of Economics and CIREQ, Université de Montréal, Quebec, Canada
[c]Department of Economics and CIREQ, McGill University, Quebec, Canada

Abstract

We compute the optimal non-linear tax policy for a dynastic economy with uninsurable risk, where generations are linked by dynastic wealth accumulation and correlated incomes. Unlike earlier studies, we take full account of the welfare distribution along the transition to the new steady state following a once-and-for-all change in the tax system. Findings show that accounting for transitional dynamics leads to a more progressive optimal tax system than one would obtain by only comparing steady states. Starting at the U.S. status quo, the optimal tax reform is a slight to moderate reduction in the progressivity of the tax system, depending on how much the policy maker cares about future generations.

Keywords: Optimal Taxation, Intergenerational Mobility, Progressive Redistribution

1. Introduction

Most modern governments implement a redistributive fiscal policy, where incomes are taxed at an increasingly higher rate, while transfers are skewed towards the poor. Such policies are thought to deliver a more equitable distribution of income and welfare, and, thereby, provide social insurance both for the currently alive, who face income fluctuations, and for future generations, who face uncertainty about what conditions they will be born

☆The authors would like to thank Arpad Abraham, Ömer Açıkgoz, Mark Bils, Rui Castro, Jonathan Heathcote, Remzi Kaygusuz, Dirk Krueger, Ananth Seshadri, Gianluca Violante, Hakkı Yazıcı and seminar participants at the NBER Summer Institute, SED, New York/Philadelphia Quantitative Macro Workshop, Federal Reserve Bank of San Francisco, University of Rochester, University of British Columbia, Sabanci University, University of Bonn, Institute for Economic Analysis Barcelona and Universitat de Barcelona for their comments.

Email addresses: bakis@sabanciuniv.edu (Ozan Bakış), baris.kaymak@umontreal.ca (Barış Kaymak), markus.poschke@mcgill.ca (Markus Poschke)

into.

In market economies, such egalitarian policies can be costly as they disrupt the efficiency of resource allocation. Therefore, the added benefit of a publicly provided social safety net, that is over and above what is available to people through other sources, such as their family or the private sector, has to be carefully weighed against this cost. In this paper, we provide such an analysis of the optimal degree of income redistribution for a utilitarian government.

The optimal design of a redistributive tax system is, however, subject to constraints. We emphasize three. First, agents may have access to insurance through other means. Savings and bequests, in particular, provide a natural source of insurance against adverse economic outcomes. A redistributive tax policy would alleviate the need for such self-insurance and crowd out accumulation of capital, leading to reduced investment.

Second, informational frictions may prevent the government from observing individual productivity. Consequently, it levies taxes on total income, which leads to well-known incentive problems as higher taxes discourage labor and thereby reduce output.

Third, the policymaker has to be cognizant of the implications of its tax policy on prices. Large-scale shifts in labor supply and savings alter the wage rate and the interest rate, which may have redistributive repercussions for income.

We explicitly address these constraints in a dynastic general equilibrium model with incomplete markets and endogenous labor supply, where generations are linked through a correlated income process. Individuals are faced with idiosyncratic fluctuations in their own income and they are uncertain about their offsprings' income. Agents do not have access to contracts contingent on future outcomes. They can, however, save and transfer wealth to subsequent generations, but may not pass their debt onto them. This is essentially an Aiyagari-Bewley-Huggett setting with a borrowing constraint.

In this setting, we search for the optimal redistributive income tax scheme. Our approach to the problem is primarily quantitative and is in the tradition of Ramsey (1927).[1] The policy maker may not modify the financial structure of the economy. It cannot, for instance, introduce new assets or allow parents to accept obligations for their kids. It may, however, implement a transfer scheme, for example to transfer income to poor agents. Transfers and government expenditures are financed by taxes levied on labor and capital income. The set of tax policies is restricted to parametric forms albeit flexible ones. The tax schedule used here not only provides a good fit to the current U.S. system, but also allows for a variety of tax systems, such as progressive, flat, and regressive taxes. We assume that the government can commit to a once-and-for-all change in the tax policy, and ask two questions: Which tax policy maximizes average welfare *at the steady state* of our model economy? Which tax policy maximizes average welfare starting from the current wealth and income distributions in the U.S., *taking into account the entire transition path* until a new steady state is reached? Since the transition to an optimal steady state may be costly, the optimal reform starting at the status quo will in general be different from the optimal steady-state policy.

We find that when the transitional dynamics are ignored, the optimal tax policy for the long-run steady state is moderately regressive. Ceteris paribus, a less progressive tax system fosters creation of wealth and income by raising the after-tax return to labor and savings, resulting in higher average consumption. The improvement in consumption levels is weighed against larger wealth and income inequality implied by regressive taxation, an undesirable feature for a utilitarian government. The latter, however, is mitigated for two reasons. First, the larger supply of capital lowers the interest rate while boosting the wage rate, as labor complements capital in production. This redistributes income away from the wealthy, who rely primarily on capital income, to consumption-poor agents who rely heavily on labor income, and counterbalances the increase in inequality generated by re-

[1] A parallel set of papers study the implications of information frictions in dynamic economies for allocations that are efficient under incentive-compatibility constraints (Mirrlees, 1971; Golosov et al., 2003; Kocherlakota, 2005; Farhi and Werning, 2012).

gressive taxation.[2] Second, the availability of self-insurance through savings considerably limits the translation of income inequality to consumption inequality. These mechanisms are effective until moderate levels of regressivity, beyond which the disutility from working even more hours outweighs that of additional income, so that hours worked do not increase further. Output and average consumption thus stop rising, while inequality keeps growing, leaving no incentive for the government to reduce progressivity further.

When the transition path is considered, a sudden switch to a regressive tax system from the current U.S. system is not desirable. Accumulation of the additional capital requires limited consumption of goods and leisure along the transition path, which limits the welfare gains from changing the tax policy. In addition, the welfare gains associated with having a higher capital stock realize only slowly since capital takes time to build. By contrast, a sudden change in the tax system involves large and immediate transfers of income which leads to substantial income inequality in the short run. Due to discounting by households, these concerns outweigh the long-run benefits of regressive income taxation.

As a result, the optimal tax reform when welfare during the transition to a new steady state is considered is much more progressive. The optimal degree of progressivity depends on how much the policy maker values future generations. When the policy maker only cares about the current generation, that is when future generations are valued only indirectly through altruistic motives of parents, a utilitarian government prefers a tax system that is close to the current status quo in the U.S.. When the policy maker values future generations directly, with the same weight that altruistic parents use, the optimal tax reform is a moderate reduction in the progressivity of the tax system.

The literature on optimal taxation is vast. The approach here is closest to Conesa and Krueger (2006) and Conesa et al. (2009), who calculate the optimal progressivity of income taxes for an OLG economy with incomplete markets and heterogeneous agents.

[2]A similar result appears in Davila et al. (2012), where saving subsidies raise the wage rate in equilibrium, and, thereby, the welfare of the poor who rely primarily on labor income.

Heathcote et al. (2014) take a similar approach to compute optimal progressivity in a Blanchard-Yaari-Bewley economy with partial insurance, and without capital. Relative to these papers, we make two contributions. First, we introduce intergenerational income risk and allow dynasties to self-insure via capital accumulation and bequests.[3] The results show that both components are important in gauging the value added by publicly provided social insurance, and for modeling the appropriate consumption response to tax policy. In particular, when self-insurance via savings is available, a benevolent government may prefer to improve social welfare by affecting the savings incentives and by harnessing general equilibrium effects rather than by directly providing insurance via income redistribution. Second, whereas the quantitative literature on optimal taxation has been limited to steady-state welfare analysis, we provide a full quantitative analysis of the optimal tax system taking into account the transition path.[4] Our findings indicate that accounting for the short-run distribution of welfare along the transition leads to a more progressive tax system than one would obtain by comparing steady states only.

Erosa and Koreshkova (2007), Seshadri and Yuki (2004) and Bénabou (2002) also look at taxation problems in dynastic settings, with emphasis on human capital investment and education. Bénabou (2002) abstracts from dynastic capital accumulation and Seshadri and Yuki (2004) from labor supply. Both Erosa and Koreshkova (2007) and Seshadri and Yuki (2004) analyze consequences of a flat tax reform, but do not calculate optimal non-linear taxation and only consider long-run stationary equilibria. Cutler and Gruber (1996), Ríos-Rull and Attanasio (2000), Golosov and Tsyvinski (2007) and Krueger and Perri (2011) study how publicly provided insurance schemes can crowd-out insurance that is available

[3] In the OLG framework of Conesa and Krueger (2006) and Conesa et al. (2009), all bequests are accidental, and they are completely redistributed among newborns. Thus, in contrast to our setting, there is neither intergenerational income risk nor self-insurance across generations in these papers. Heathcote et al. (2014) abstract from capital altogether for tractability reasons.

[4] A notable exception is Domeij and Heathcote (2004), who account for welfare along the transition in their analysis of linear capital taxes. More recently, Fehr and Kindermann (2014) compute optimal income tax progressivity in an incomplete markets OLG economy à la Conesa et al. (2009), and Krueger and Ludwig (2013) analyze optimal progressivity in a similar setting with human capital accumulation.

through other sources. Hubbard et al. (1995), in particular, emphasize the crowding out of precautionary savings by public tax policy.

In what follows, we introduce the model and formally define the optimal taxation problem. Section 3 describes our calibration. Sections 4 presents the optimal tax policy based on the comparison of steady-state economies and 5 presents the optimal tax policy along the transition. Section 6 concludes.

2. A Dynastic Model with Redistributive Income Taxation

The model is a standard model of savings with uninsured idiosyncratic income risk (Aiyagari, 1994; Bewley, 1986; Huggett, 1993) extended to incorporate intergenerational dynamics, non-linear fiscal policy and endogenous labor supply.

The economy consists of a continuum of measure one of heterogeneous consumers, a representative firm, and a government. Each consumer is endowed with capital, k, and a stochastic labor skill, z. With these endowments, they can generate an income of $y = zwh + rk$, where w is the market wage per skill unit, $h \in (0, 1)$ is hours worked and r is the net real interest rate. Each period, every consumer faces a probability μ of dying and being replaced by a descendant, who inherits her savings. The intergenerational transmission of z is described below.

Agents pay taxes on their income to finance an exogenous stream of government expenditure, g_t. The disposable income of an agent net of taxes is given by the function $y^d(y)$, which depends only on the agent's total income. This function also determines the distribution of the tax burden.

Agents allocate their disposable income between consumption and capital investment to maximize the expected present value of their utility. They derive utility from consumption, and they dislike work. In addition, agents care about their offspring's welfare, which depends on the amount of wealth passed on by the parent, and the child's skill endowment. Hence, agents save both to insure against fluctuations in labor efficiency over their life and to transfer wealth to their offspring when they die. They are not, however, allowed to borrow. We use z to denote an agent's labor efficiency units. The fluctuations in z over the life cycle and across generations are captured by a first-order Markov process: $F(z'|z)$. We describe this process in detail below.

The problem of an agent is to choose labor hours, consumption and capital investment to maximize the expected present value of the dynasty's utility. The wage rate, the interest rate and the aggregate distribution of agents over wealth and productivity, denoted by Γ, are given. Let $\Gamma' = H(\Gamma)$ describe the evolution of the distribution over time. The Bellman equation for a consumer's problem then is:

$$V(k, z; \Gamma) = \max_{c, k' \geq 0,\ h \in (0,1)} \left\{ \frac{c^{1-\sigma}}{1-\sigma} - \theta \frac{h^{1+\epsilon}}{1+\epsilon} + \beta \mathbb{E}[V(k', z'; \Gamma')|z] \right\} \tag{1}$$

subject to

$$\begin{aligned} c + k' &= y^d(y) + k \\ \Gamma' &= H(\Gamma). \end{aligned}$$

The production technology of a representative firm uses aggregate capital, K, and labor, N, as inputs, and takes the Cobb-Douglas form: $F(K, N) = K^{\alpha} N^{1-\alpha}$. Factor markets are competitive, and firms are profit maximizers.

A *competitive equilibrium* of the model economy consists of a value function, $V(k, z; \Gamma)$, factor supplies, $k'(k, z; \Gamma)$ and $h(k, z; \Gamma)$, a wage rate, $w(\Gamma)$, an interest rate $r(\Gamma)$, and an evolution function $H(\Gamma)$ such that:

(i) Given $w(\Gamma)$, $r(\Gamma)$ and $H(\Gamma)$, $V(k, z; \Gamma)$ solves the worker's problem defined by (1) with the associated factor supplies $k'(k, z; \Gamma)$ and $h(k, z; \Gamma)$.

(ii) Factor demands are given by the following inverse equations:

$$r(\Gamma) = \alpha(K/N)^{\alpha-1} - \delta$$
$$w(\Gamma) = (1-\alpha)(K/N)^{1-\alpha}$$

(iii) Markets clear:

$$K' = \int k'(k,z)d\Gamma(k,z) \quad \text{and} \quad N = \int zh(k,z)d\Gamma(k,z).$$

(iv) $H(\Gamma)$ is consistent with $F(z'|z)$ and the savings policy $k'(k,z;\Gamma)$.

(v) The government budget is balanced:

$$g = \int [y - y^d(y)]d\Gamma(k,z).$$

A *steady state* of the economy is a competitive equilibrium where the distribution of agents is stationary, i.e. $\Gamma^{ss} = H(\Gamma^{ss})$.

2.1. A Redistributive Income Tax Policy

Taxes are modeled after the current U.S. income tax system. Following Bénabou (2002) and Heathcote et al. (2014), we approximate disposable income with a log-linear function in gross-income.

$$y^d = \lambda(zwh + rk)^{1-\tau}. \tag{2}$$

The power parameter $\tau \leq 1$ controls the degree of progressivity of the tax system, while λ adjusts to meet the government's budget requirement. When $\tau = 0$, the equation above reduces to the familiar proportional tax (or flat tax) system. When $\tau = 1$, all income is pooled and redistributed equally among agents. For $0 < \tau < 1$, the tax system is progressive.[5]

The disposable income function above also allows for negative taxes. Income transfers are, however, non-monotonic in income. When taxes are progressive, transfers are first increasing, and then decreasing in income. Examples of such transfers schemes include the earned income tax credit, welfare-to-work programs etc. In Section 3, we show that this functional form provides a remarkable fit to the U.S. tax system.

A regressive tax system is achieved when τ is negative. In this case, taxes are first increasing, then decreasing in income for high enough income levels, and may prescribe positive transfers for high income earners. Since the marginal tax rate, $1-\lambda(1-\tau)y^{-\tau}$, is monotonic in pre-tax income, (2) rules out tax policies that are progressive for some parts of the income distribution and regressive elsewhere.

2.2. Intergenerational and Life-Cycle Dynamics

The process for labor efficiency units is modeled as:

$$\ln z_{igt} = m_z + f_{ig} + a_{it}, \tag{3}$$

where i indexes dynasties, g generations and t time. $f_{ig} \in \{f_L, f_H\}$ denotes the intergenerational component of productivity, which remains fixed during each individual's life, and $a_{it} \in \{a_L, a_H\}$ denotes the life cycle component of productivity, which may change from period to period. Let F and A be the transition matrices for these components. If an agent survives to the next period, which happens with probability $1-\mu$, his labor efficiency is

[5]The average tax rate is $1-\lambda y^{-\tau}$, which is increasing in y if $\tau > 0$.

determined by the transition matrix

$$S = \left(\begin{array}{c|cccc} & f_L + a_L & f_L + a_H & f_H + a_L & f_H + a_H \\ \hline f_L + a_L & A_{11} & A_{12} & 0 & 0 \\ f_L + a_H & A_{21} & A_{22} & 0 & 0 \\ f_H + a_L & 0 & 0 & A_{11} & A_{12} \\ f_H + a_H & 0 & 0 & A_{21} & A_{22} \end{array} \right)$$

Since f_{ig} if fixed over an agent's life, S is block-diagonal. If instead an agent dies, which occurs with probability μ, his offspring's productivity is determined by the matrix

$$D = \left(\begin{array}{c|cccc} & f_L + a_L & f_L + a_H & f_H + a_L & f_H + a_H \\ \hline f_L + a_L & \pi F_{11} & (1-\pi)F_{11} & \pi F_{12} & (1-\pi)F_{12} \\ f_L + a_H & \pi F_{11} & (1-\pi)F_{11} & \pi F_{12} & (1-\pi)F_{12} \\ f_H + a_L & \pi F_{21} & (1-\pi)F_{21} & \pi F_{22} & (1-\pi)F_{22} \\ f_H + a_H & \pi F_{21} & (1-\pi)F_{21} & \pi F_{22} & (1-\pi)F_{22} \end{array} \right)$$

where π denotes the share of newly born agents who start their careers with a_L. When π is high enough, agents, on average, start their career with lower productivity, and tend to improve later as their careers progress. This helps generate wage growth over the life-cycle.[6] When F_{11} and F_{22} are greater than a half, a positive correlation of wages emerges across generations. Note that the intergenerational transition probabilities are independent of the level of the life-cycle component, a_{it}, at the time an agent dies. This enables us to estimate the elements of the intergenerational transition matrix based on the permanent component of wages as is standard in the empirical literature.

The aggregate transitions across different endowments of labor efficiency units in the

[6]For this to happen, π has to be greater than $A_{21}/(A_{21} + A_{12})$, the share of workers with a_L at the stationary distribution associated with A.

economy depend on both life-cycle and intergenerational transitions and are given by $\mu D + (1-\mu)S$.

2.3. Optimal Taxation Problem

The benevolent Ramsey government maximizes average welfare in the economy by choosing the progressivity of the tax policy subject to a balanced budget constraint and equilibrium responses by households to the tax policy. We conduct a series of experiments, using different objective functions for the policy maker. In the first experiment, the policy maker is concerned with the average welfare at the long-run steady state of the economy. Formally, the problem is:

$$\max_{\lambda,\tau} \mathcal{W}^{ss} = \int V^{ss}(k,z;\Gamma^{ss}) d\Gamma^{ss}(k,z)$$

subject to

$$g = \int [y - y^d(y;\lambda,\tau)] d\Gamma^{ss}(k,z) \quad (4)$$

$$y = wzh(k,z;\Gamma^{ss}) + rk'(k^-,z^-;\Gamma^{ss}). \quad (5)$$

where V^{ss} is the value function, Γ^{ss} is the stationary distribution of agents over productivity and wealth, $h(.)$ and $k'(.)$ are the policy functions at the steady-state equilibrium associated with the tax policy (λ,τ), and k^-, z^- are the lagged values of k and z such that $k'(k^-,z^-) = k$. The dependence of these functions on the tax policy is suppressed for notational convenience. The steady-state experiment turns out to be very useful for understanding the tradeoffs the policy maker faces.

However, the steady-state objective ignores welfare along the transition to the new steady state. In the remaining experiments, we therefore assume that the economy starts off in the status quo, and that the government can credibly commit to a *once-and-for-all* change in the tax policy. We assume that the tax reform takes effect in period 0 and is not anticipated. In the second policy experiment, the policy maker seeks to maximize average utility by choosing the parameters of the tax reform. Let $\tilde{\Gamma}_t(k, z)$ be the distribution of agents born in period t.[7] Formally, the policy maker solves

$$\max_{\lambda,\tau} \int V_0(k, z; \Gamma_0) d\Gamma_0(k, z) + \mu \sum_{t=1}^{\infty} \beta_g^t \int V_t(k, z; \Gamma_t) d\tilde{\Gamma}_t(k, z) \tag{6}$$

subject to a balanced budget constraint each period and optimal, competitive behavior on part of the agents. The first component is the average welfare of agents that are alive when the new tax policy is implemented. The second component is the welfare of future generations, discounted at a rate β_g by the policy maker. μ is the total measure of newborn agents that enter the economy each period. When $\beta_g = 0$, the policy maker only takes into account the welfare of those who are alive in the initial period, at the time of the reform. Future generations appear in the policy maker's objective function only indirectly, due to parental altruism of the existing generations. When $\beta_g > 0$, future generations are valued both directly, and indirectly through their parents' welfare (see Farhi and Werning (2007), for instance, for a discussion). Below, we report results for different values of β_g.[8,9]

Due to the concavity of the individual utility function, the utilitarian welfare criterion favors redistribution even when there are no shocks to be insured. It therefore confounds

[7] Since newborns start their career with a lower wage rate (a_L), their distribution over the z space is different from Γ_t. Specifically, if $P(z)$ denotes the stationary measure of agents with labor efficiency z, and $\tilde{P}(z)$ the measure of newborns with z, $\tilde{\Gamma}(k, z) = \Gamma(k, z)\tilde{P}(z)/P(z)$ for all k.

[8] Note that the objective function defined by (6) differs from the typical Ramsey problem as formulated by Chamley (1986) or Aiyagari (1995). These papers allow for time-varying tax rates and theoretically study their limiting values at the long-run stationary state equilibrium. By contrast, we fix the tax rate in the initial period and over the transition. Recently, Acikgoz (2013) illustrates that in an Aiyagari-type model, long-run optimal fiscal policy can be computed without explicitly solving for the optimal transition. To our knowledge, the computation of time-varying tax rates during the transition in the Aiyagari-Bewley-Huggett type of models remains an open question.

[9] There are two main differences between the welfare function defined in (6) and that used in comparable studies in an OLG setting (Conesa et al., 2009). First, our formulation includes the welfare of existing generations in period 0, and does not focus on newborns only. Second, our newborns have a non-degenerate asset distribution due to bequests, whereas all newborns in Conesa et al. (2009) start with zero assets. See Fehr and Kindermann (2014) for a discussion of welfare functions in that context.

the insurance motive of redistribution with the pure equity motive. To separate the efficiency concerns from equity concerns, we conduct another experiment and employ a version of the aggregate efficiency criterion introduced by Bénabou (2002). The idea is to replace the stochastic consumption sequence of an agent (and his dynasty) with its certainty-equivalent, and to evaluate welfare by aggregating the certainty-equivalent levels of consumption rather than utility levels in each period. With this procedure, risk aversion is reflected in the certainty-equivalent evaluation of consumption streams, but not in the interpersonal aggregation. As a consequence, redistribution then matters insofar as it provides insurance, reduces risk, and changes certainty-equivalent consumption, but is not valued for interpersonal redistribution.

When the utility function depends not only on consumption but also leisure, there are multiple ways to compute certainty equivalence depending on how disutility from work is treated. Since the utility function is additively separable between consumption and hours worked, our approach is to compute the certainty equivalence for each component separately. Formally, let $V_0(k,z) = V_0^c(k,z) - V_0^n(k,z)$ be the value of initially having a capital stock of k and labor efficiency z. The certainty equivalent levels of consumption and hours, denoted by $\tilde{c}(k,z)$ and $\tilde{n}(k,z)$, solve the following set of equations.

$$V_0^c(k,z) = \mathbb{E}\sum_{t=0}^{\infty}\beta^t \frac{c(k_t,z_t)^{1-\sigma}}{1-\sigma} = \frac{1}{1-\beta}\frac{\tilde{c}(k,z)^{1-\sigma}}{1-\sigma}$$
$$V_0^l(k,z) = \theta\mathbb{E}\sum_{t=0}^{\infty}\beta^t \frac{h(k_t,z_t)^{1+\epsilon}}{1+\epsilon} = \frac{1}{1-\beta}\frac{\tilde{n}(k,z)^{1+\epsilon}}{1+\epsilon}$$

Finally, the associated objective function is defined as follows.

$$\mathcal{W}^E = \frac{1}{1-\sigma}\left(\int \tilde{c}(k,z)d\Gamma_0(k,z)\right)^{1-\sigma} - \frac{\theta}{1+\epsilon}\left(\int \tilde{n}(k,z)d\Gamma_0(k,z)\right)^{1+\epsilon}. \qquad (7)$$

The purpose of this experiment is not to dismiss equity concerns, but rather to be able to separately evaluate equity and efficiency concerns, which in the utilitarian welfare function are simultaneously present. A comparison of results for this objective with those for the utilitarian objective thus allows assessing how optimal progressivity is affected by equity concerns, and what it would be with efficiency concerns only. We also acknowledge that alternative formulations of the efficiency criterion may yield slightly different results.

3. Empirical Analysis and Calibration

The model is calibrated to the U.S. economy. For computational convenience, the model period is set to 5 years. μ is set to 0.2, which implies that in expectation, each generation of a dynasty holds the dynasty's capital for 25 years. The capital share of income, α, is set to 0.36, the depreciation rate to 8% per annum and the rate of relative risk aversion to 2. This leaves three sets of parameters: the fiscal policy, (g, λ, τ), the preference parameters for labor, θ and ϵ, and the parameters for the stochastic income process, z and $F(z'|z)$. These parameters are identified as follows.

3.1. How Progressive *is* the U.S. Tax System?

The progressivity of the current tax system is estimated using household-level data from March supplements to the Current Population Survey for 1979 to 2009. Federal and state income taxes, as well as the payroll tax per household are obtained from the NBER tax simulator (Feenberg and Coutts, 1993). Our measure of pre-tax income is gross earnings, as reported by the household, plus the payroll tax. Disposable income is defined as reported earnings less federal and state income taxes. The estimated log-linear regression is:

$$\log y^d(y) = 1.34 + 0.83 \log y + X\hat{\Gamma} \qquad R^2 = 0.94. \tag{8}$$

To control for the changes in the average tax rate over the years, X includes indicator variables for each survey year.[10] The correlation coefficient indicates that the log-linear specification fits the U.S. tax system remarkably well. Figure 1 further confirms this visually by plotting average disposable income by quantiles of pre-tax income (circles) over the regression line (solid). Two points are worth noting. First, the slope of the regression line is less than one, showing the progressivity of the U.S. tax system. The implied value of τ is 0.17 (0.0026).[11] Second, the bottom five percent of the gross-income distribution are paying negative or zero taxes.

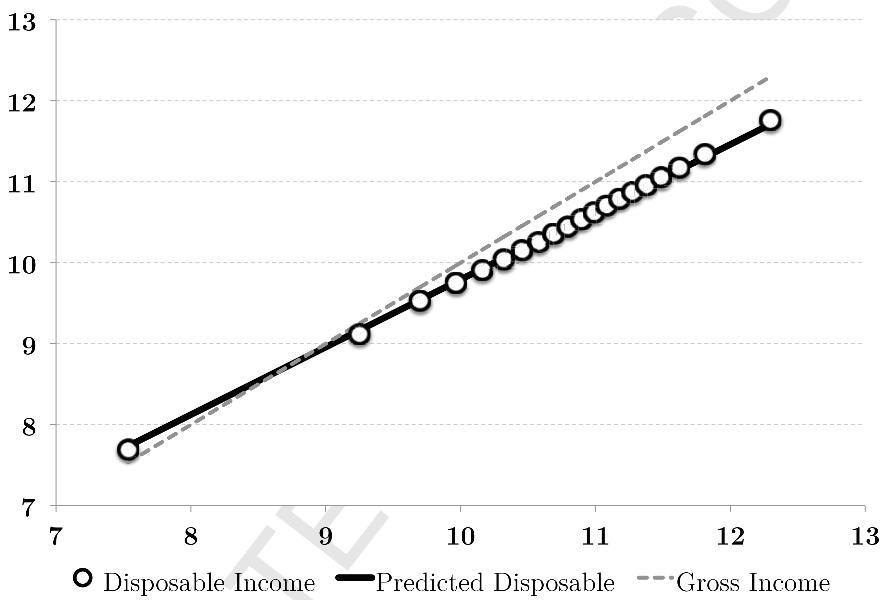


Figure 1: The progressivity of the U.S. tax system – Disposable household income as a function of pre-tax income. Circles denote average income net of taxes by quantiles of gross income. The solid line is the fitted regression line and the broken line is the 45 degree line. Data combines March supplements to CPS (1979 - 2009) with the NBER tax simulator.

As Conesa et al. (2009), we set government expenditures to 17% of output in the benchmark economy. The implied *level* of government expenditure is then kept constant

[10]Failing to control for year effects leads to a downward bias in the estimated tax progressivity, $\hat{\tau}$, since G/Y is countercyclical.

[11]Corporate taxes are not available in our dataset. To test the relevance of this for our estimate, we estimated the same specification for 2004 based on the information in Table 2 of Piketty and Saez (2007), who impute corporate taxes in their calculations using federal tax returns. We estimate the progressivity to be 0.164, virtually the same as our estimate above. Guner et al. (2014) reports a lower estimate using federal tax data but excluding transfers. Heathcote et al. (2014) report an estimate of 0.185 when time effects are excluded from the regression.

when evaluating alternative tax policies. Given $\tau = 0.17$ and $g/Y = 0.17$, the value of λ is determined at the equilibrium by the government's budget constraint.

3.2. Income Process and Intergenerational Dynamics

To calibrate the components of the transition matrix, we estimate the transition probabilities in A and F using panel data on hourly wages from the PSID (1970 - 1991). The sample is restricted to men of ages 24 to 60 who report to be household heads. To decompose the wages into its life cycle and intergenerational components, we estimate the following specification:

$$\ln w_{it} = \phi_{ig} + g(age_{it}; \Phi) + I_t + \varepsilon_{it}, \tag{9}$$

where ϕ_{ig} denotes the fixed effect for worker i of generation g. Since fathers and sons may be observed at different points in the life-cycle, and possibly at different points of a business cycle, indicators for survey year, I_t, and a quartic polynomial in age, $g(age_{it}; \Phi)$, are included as control variables.

The intergenerational component of wages is defined by $f_{ig} = \hat{\phi}_{ig}$ and the life-cycle component by $a_{it} = g(age_{it}; \hat{\Phi}) + \hat{\varepsilon}_{it}$, where the first terms captures the the deterministic age profile of wages, and the second term captures the transitory shock to the wage rate. The variance of the intergenerational component in the data is 0.22 and that of the life cycle component is 0.12, implying that the fixed worker effects capture 64% of the total wage variance. This is in line with the findings in Storesletten et al. (2004) who estimate the share of fixed worker effect to be around 56% for earnings.

There is a longstanding literature on the intergenerational income mobility in US (see Solon (1999) for a survey). The elasticity of offspring's earnings to parental earnings reported in the literature is around 0.40. The elasticity of wages is usually slightly lower than the earnings elasticity due to the positive intergenerational correlation of hours. Using data from the PSID, Solon (1992) reports an intergenerational wage elasticity of 0.30.

Similarly, Mulligan (1997) reports a wage elasticity of 0.33. The intergenerational wage elasticity in our sample is 0.32.

Our model allows for two values of f: f_L and f_H. To estimate the intergenerational transition probabilities between these levels, we split the distributions of f_{ig-1} (fathers) and f_{ig} (sons) into two quantiles around the median, and compare the son's position relative to the median in the f_{ig} distribution given his father's position in the f_{ig-1} distribution.[12] Similarly life-cycle transition probabilities are estimated by splitting the distributions of a_{it} and a_{it-1} into two quantiles, and tracking a worker's position relative to the median in two consecutive years.

The corresponding states (a_L, a_H) and (f_L, f_H) are calibrated such that the means of a_{it} and f_{ig} at the stationary state are (normalized to) zero and the standard deviations of each component match their data counterparts. The resulting states are $(f_L, f_H) = (-0.40, +0.55)$, and $(a_L, a_H) = (-0.37, +0.33)$. Combined with the average wages, these states imply four possible values for hourly wages: \$8.3, \$16.7, \$21.5 and \$43.2 in 1999 dollars.[13]

The transition matrices corresponding to each component are below.[14]

$$F = \begin{bmatrix} 0.69 & 0.31 \\ 0.43 & 0.57 \end{bmatrix} \qquad A = \begin{bmatrix} 0.68 & 0.32 \\ 0.28 & 0.72 \end{bmatrix}.$$

Note that the matrices are not symmetric. For intergenerational transitions, the asymmetry stems from limited upward mobility across generations in the U.S., an aspect well documented in the literature (Solon, 1999; Black and Devereux, 2011). In the case of life cycle transitions, the asymmetry comes from the wage growth over a worker's career.

[12] Following the literature, we replicate the wage observation for the father in the case of fathers with multiple sons.

[13] All income values were adjusted by the CPI.

[14] Matrix A is reported at an annual frequency. The corresponding persistence for the transitory component, a_{it}, is 0.44. Combined with the fixed component, annual persistence of wages in the data is 0.78.

Wage growth over the life cycle depends not only on the transmission probabilities in A, but also the initial distribution of new entrants over a_L and a_H. Therefore, π is calibrated to observed wage growth over the life cycle given the probabilities in A. In the data, wages grow rapidly during the first 20 years of a worker's career and remain flat thereafter, until the last 5 years of the career, when wages decline slightly before retirement. Therefore, we target wage growth during the first 20 years (4 model periods) to determine π. The average log-wage difference between workers aged 44 to 49 and those aged 24-29 in the data is 0.30. Given the estimated transition matrices above, this is replicated when $\pi = 0.9$, implying that 90% of all newborn agents starts their career with a_L.

3.3. Leisure and Labor Supply

The discount factor β, the preference parameter for labor disutility, θ, and the curvature of utility with respect to hours worked, ϵ, are jointly calibrated to an annual interest rate of 4.1%, average hours worked over life, and the coefficient of variation of average lifetime labor hours. To obtain the latter components in the data, we estimate the specification in (9) for annual working hours in a year. An average person works 2,122 hours in the data. We consider this to be approximately 49% of available work time during the year.[15] The standard deviation of the estimated fixed worker effects for hours implies a coefficient of variation of 0.27 for hours worked.

Table 1 summarizes the calibrated values for the parameters. The implied values for the utility parameters are: $\theta = 0.358$, $\epsilon = 1.183$, and $\beta = 0.962^5$. The Frisch elasticity in the model depends both on the utility parameter ϵ, and the progressivity of the tax system τ as follows: $(1 - \tau_l)/(\epsilon + \tau_l)$. The calibrated values for these parameters imply an elasticity of 0.62. This is somewhat higher than the micro level estimates for yearly models, which are around 0.25 for individuals, while a value between 2 and 3 is required to match

[15] Total available time in a year is considered to be 52 weeks of 7 days each, with 12 hours available for work each day.

Table 1: Calibration of the Model to the U.S. Economy

Parameter Value		*Target Moment*	
σ	= 2.00	relative risk aversion	2.00
$\beta^{1/T}$	= 0.962	annual interest rate (McGrattan and Prescott, 2010)	4.1%
θ	= 0.358	average annual labor hours	0.49
ϵ	= 1.183	coef. of variation of hours	0.27
α	= 0.36	capital share of income	0.36
δ	= 0.34	annual depreciation rate	0.08
g	= 2.59	g/Y (Conesa et al., 2009)	0.17
τ	= 0.17	own estimate (see Section 3.1)	

employment differences across time and countries at the macro level (See Prescott (2004); Cho and Cooley (1994) and Blundell and MaCurdy (1999) among others). More recently, Blundell et al. (2012) report an estimate of 0.4 for males and 0.6 for females. Thus a value of 0.62 for a model that does not distinguish individuals by gender seems broadly plausible.

The calibration ensures that the benchmark economy exactly delivers the wage inequality in the PSID (1970-1991). Combining this with the labor supply policies results in a Gini coefficient for earnings of 0.35, which is closely in line with the data for this period (Heathcote et al., 2010). The model thus replicates wage dispersion, inter- and intragenerational wage dynamics, and earnings inequality in the PSID very well. The model does slightly less well in terms of the wealth distribution. The Gini coefficient for wealth in the model is 0.53, which is below the reported estimates in the data. The data used to estimate the wage process spans the years between 1970 to 1991. Heathcote et al. (2010) reports a wealth Gini of 0.66 during this period based on the Survey of Consumer Finances. More recently, the wealth Gini has exceeded 0.8 (Diaz-Gimenez et al., 2011). As a consequence, inequality in capital income and thus in total income is also somewhat below data values. This issue is not unique to our model; see e.g. Heathcote et al. (2010, p. 41), who state that "one should not expect a model calibrated to wage or income dynamics from the PSID to replicate the extreme wealth inequality in the raw SCF."[16] Importantly, while our model mostly understates wealth inequality in the upper tail, it does much better for the lower tail. In particular, it captures the fact that the bottom 10% of households hold no net wealth at all (Heathcote et al., 2010). Matching this is important, as these households are constrained and thus are strongly affected by redistributive policy. Since these agents have low consumption levels, they also carry the highest welfare weight for a utilitarian policy maker.

4. How Progressive Should the Long-Run Tax Policy Be?

In this section, we report results for our first experiment by comparing the steady-state equilibria under different tax policies with varying degrees of progressivity. The findings suggest that the optimal tax code in the long run is slightly more *regressive* than a flat tax system. The optimal value for τ is -0.09. Below, we highlight several key aspects influencing the policy maker's decision. This will also help understand the trade-off between the long-run welfare gains and short-run costs incurred along the transition analyzed in the next section. These costs are instrumental to the optimality of *progressive* income taxes when transition dynamics are considered.

How could a regressive tax system, which subjects low income groups to higher tax rates, be optimal for an egalitarian government? To see this, note that a utilitarian policymaker is concerned with two things when comparing tax policies: the total amount of available goods (consumption and leisure), and how these goods are distributed among agents. A less progressive tax policy raises average consumption at the cost of higher *after-tax* income inequality. However, this does not translate to equally severe consump-

[16]Modeling devices that help to match wealth inequality are the presence of entrepreneurs, who have a strong incentive to accumulate wealth as their private return exceeds the market return (Cagetti and De Nardi, 2006), or the presence of "superstars" and retirees (Castaneda et al., 2003). Information on the intergenerational transmission of these traits is limited in the PSID.

Table 2: Optimal Tax System: Steady-State Comparison

	US	Optimal		US		Optimal	
Progressivity (τ)	0.17	-0.09	Output	15.2		+18.4%	
Interest Rate (%)	4.10	2.60	Pre-Tax Income	12.5	(0.30)	+34.4%	(0.32)
Wage Rate	0.50	0.55	Disposable Income	9.4	(0.25)	+34.0%	(0.35)
Hours	0.49	0.52	Wealth	9.7	(0.53)	+40.2%	(0.56)
G/Y	0.17	0.14	Consumption	9.3	(0.17)	+16.1%	(0.20)

Note.– Table compares the benchmark US economy with the optimal tax system that maximizes average long-run welfare at the steady state. The numbers in parentheses show the Gini coefficients of inequality.

tion inequality when agents self-insure with precautionary savings. As a consequence, the optimal tax schedule may well be regressive.

Table 2 compares the steady state of the benchmark economy calibrated to the U.S. tax policy with that obtained under the optimal tax system. A decline in the progressivity of the tax policy promotes generation of income by increasing the after-tax return to labor and capital. This raises savings in the economy. Lack of redistribution also leads to a higher income risk, and promotes additional precautionary savings. For high-income groups, there is a positive income effect generated by lower taxes, which further encourages savings. For low-income groups, the income effect works against the substitution effect, but is not strong enough. Overall, the supply of capital increases, which puts a downward pressure on the interest rates.

The larger capital stock has two implications for labor. First, it raises the demand for labor, and increases the wage rate, despite the downward pressure created by the increase in the labor supply. Second, larger wealth has a negative income effect on labor supply, limiting the increase in labor input, and pushing the wage rate further up. With a larger stock of capital and increased labor input, output increases. The optimal tax system leads to a 18.4% increase in output, which translates to a 16.1% increase in consumption. The rise in welfare due to higher average consumption is mitigated by the decrease in average leisure from 0.51 to 0.48. Since government expenditure g is fixed in all our experiments, higher output implies a lower average tax burden, and g/y falls from 0.17 to 0.14.

Overall, an average person has larger wealth, higher income, substantially more consumption, and less leisure. To compare this improvement in the utility of an average person with the change in distributive inequality, the Gini coefficients of inequality are shown in parentheses in Table 2. Pre-tax income inequality remains similar under regressive taxes due to the change in the equilibrium prices. The decline in the interest rate attenuates the effect of wealth inequality on income, while the higher wage rate raises labor income, which is distributed more equally. Nonetheless, the economy with regressive taxes features larger wealth inequality along with a considerable increase in the inequality of after-tax income disposable for consumption. The Gini coefficient for wealth inequality increases from 0.53 to 0.56, and that for disposable income from 0.25 to 0.35. The impact of rising income and wealth inequality on consumption, however, is limited. The Gini coefficient for consumption inequality rises from 0.17 to 0.20, about a third of the rise in disposable income inequality. This is due, in large part, to the availability of self-insurance through precautionary savings.

4.1. Tax Progressivity and Steady-State Welfare

To gauge the improvement in average steady-state welfare, we ask the following hypothetical question: by what factor would one need to increase the consumption of each and every person in the benchmark economy to reach the same average welfare as the optimal economy, keeping their labor supply constant? The answer is 3.3%, which is quite large considering that the welfare cost of business cycles are estimated at 1% or less, even for models with heterogeneous households.[17] This calculation ignores the change in welfare during the transition to the new steady state, which is studied in Section 5.

To see how the welfare distribution changes across agents, consider first the value func-

[17] For a risk aversion of two (as here), Krebs (2007) reports 0.98%, which is much larger than the estimate reported in Lucas (1987).

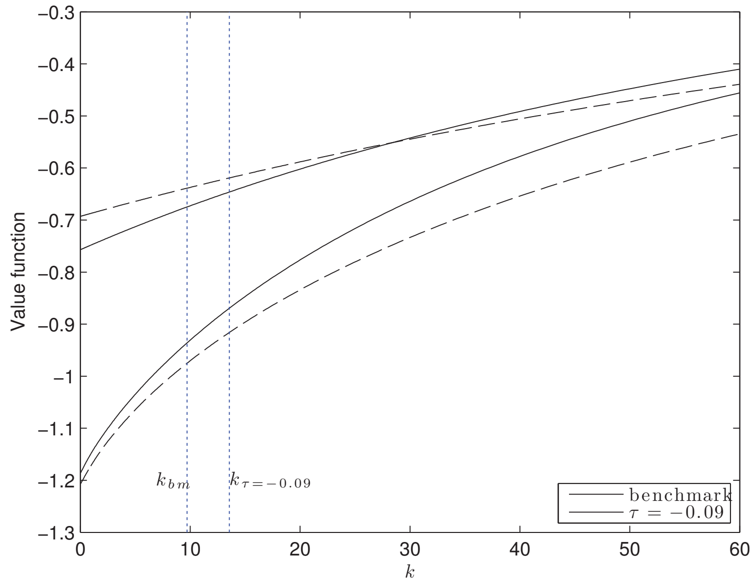


Figure 2: Welfare by Wealth and Productivity – Note: the vertical lines indicate average wealth in the two economies.

tions for a given wealth and productivity level, without taking into account the shift in the wealth distribution. Figure 2 plots welfare by wealth for the lowest and the highest productivity groups (out of 4 in total). The solid lines correspond to the benchmark economy, and the dashed lines represent the economy operating under the optimal tax code. The optimal economy features lower welfare for the wealthy, especially for those with little labor income. This is primarily due to the lower interest rate in the optimal economy. Workers with low wealth, on the other hand, are dependent on labor income, which is higher in the new economy due to higher wage rates. This leads to higher welfare for the highly productive, who have higher disposable incomes in the new tax system, and eases the fall in welfare for workers with low productivity and, hence, low income, who receive less transfers under the regressive tax system.

A utilitarian policymaker also considers the shift in the wealth and income distributions when comparing these two economies. In particular, the optimal economy features a higher wealth level on average, which leads to an upward movement along the dashed welfare functions in Figure 2.

A similar intuition emerges in Davila et al. (2012), who show that in Aiyagari (1994) type models, agents' individual savings decision in the laissez-faire equilibrium imply that the economy does not reach its constrained efficient optimum. They show that in a situation where the income of the poor consists mainly of labor income, as also is the case here, a subsidy to saving with the objective of promoting capital accumulation improves the welfare of the poor by raising the wage rate. Reinstating constrained efficiency in that paper requires state-dependent tax and transfer schemes on capital income at the individual level. Our findings suggest that when such policies are not feasible, a regressive income tax system can stand in for a more complicated mechanisms.[18]

4.2. Labor Supply, Self-Insurance and Partial Equilibrium: Implications for Tax Policy

In this section, we conduct counterfactual experiments to highlight the roles of the three constraints on the policymaker's choice of redistributive tax policy: the crowding out of labor supply, the crowding out of self-insurance and adjustment of prices in equilibrium.

First, we recompute the optimal tax code assuming that savings behavior remains fixed at the benchmark economy. Agents are allowed to optimally adjust their labor supply, prices clear markets, and the budget is balanced at all times. Since savings are fixed, lower progressivity leads to larger consumption inequality with little improvement in aggregate output or consumption. Consequently, the optimal tax policy is highly progressive with a τ of 0.36.

To gauge the role of equilibrium price adjustments, the interest rate and the wage rate

[18]These results seem to contrast with Aiyagari (1995), who shows that the optimal long-run tax on capital is positive in an incomplete markets setting due to 'excessive' saving arising from the precautionary savings motive. Three crucial ingredients to his finding are absent here: inelastic labor supply, time-varying tax rates and the presence of government debt. Therefore, his result is not directly comparable to ours. As shown by Marcet et al. (2007), already the presence of elastic labor supply implies that there is no 'excessive' saving in Aiyagari–Bewley-Huggett type models if the income elasticity of hours worked is large relative to the income elasticity of consumption.

are fixed at their benchmark levels in the second experiment. Savings and labor supply respond optimally, and the government runs a balanced budget. In this partial equilibrium exercise, the redistributive role of prices in response to lower progressivity is absent. As a result, the optimal tax system is progressive with a τ of 0.23.

Finally, when the labor supply response is shut down, the optimal tax system remains regressive with $\tau = -0.06$. This is because the long-run elasticity of labor supply is low due to two contradicting effects. On the one hand, progressive taxes reduce labor supply by lowering the return to an hour of work. On the other hand, they reduce the steady-state income in the economy, creating a negative income effect on leisure.

These experiments reveal that two features are key for the optimality of a tax system when only long-run outcomes are considered: the effect of taxes on saving, and the resulting changes in equilibrium prices. These features are absent in settings where there is no effect of capital accumulation on wages (as in Heathcote et al., 2014), or where assets cannot be transmitted to the next generation, limiting the response of savings to the tax system (as in Conesa and Krueger 2006 and Conesa et al. 2009). In both cases, public insurance becomes more attractive.

5. Optimal Redistribution along a Transition Path

The optimal tax code described in the previous section encourages capital accumulation and accordingly leads to higher wages than the benchmark economy. Getting there is costly, however, as building new capital requires initially reducing consumption and leisure. Therefore, the transition to the steady state following a switch to a regressive tax system is costly in terms of welfare. Comparing steady states abstracts from this cost. Depending on its size, implementing the tax code that is optimal in the long run may not be optimal once the transition is taken into account. Overall welfare including the transition may instead be maximized by a completely different tax code. Therefore, we next ask the following two questions: What are the short-run implications of implementing the tax code that is optimal at the steady state of the economy? And which level of progressivity of the tax code is optimal, taking into account the transition from the current U.S. benchmark?

5.1. Transition to the Optimal Steady State

To analyze the transitional dynamics, we assume that the economy is initially in the benchmark steady state that reproduces the U.S. status quo. In this situation, the government surprisingly implements the new tax code and commits to it. As the economy converges to the new steady state under the tax system, the interest rate, the wage rate and λ all change. Recall that the parameter λ of the tax code adjusts to balance the government's budget every period.

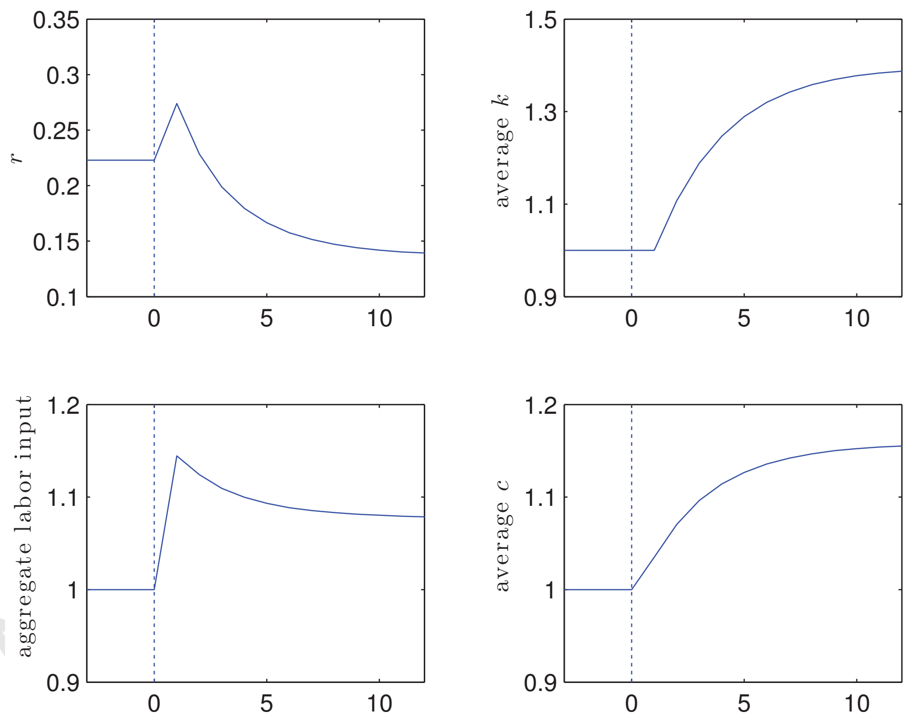


Figure 3: Transition to Regressive Taxation – Values relative to the benchmark economy, except for r. The new tax system with $\tau = -0.09$ goes into effect at $t = 0$.

Results show that the transition to the optimal long run-policy is very costly. Values of key endogenous variables along the transition path are shown in Figure 3. The economy moves into the neighborhood of the new steady state in 25 to 50 years. Over this time, the capital stock increases by more than a third and average consumption rises by about 15%. Early in the transition, however, increased capital accumulation requires a 14% increase in labor hours, implying a significant reduction in leisure. Furthermore, a sudden change in the tax policy brings about a substantial increase in the after-tax income inequality, and, thereby, consumption inequality. The Gini coefficient for consumption rises from 0.17 to 0.20 immediately and remains relatively stable thereon. Unlike the rise in consumption inequality, average consumption rises only by 3.5% in the first period, approximately a fifth of the overall increase in consumption in the long run.

Due to discounting, the short-run welfare costs associated with switching to the regressive tax system carry a higher weight than long term gains. Using the same method as in Section 4.2, consumption of each person in the current generation would need to be *reduced* by 3% at every time and state in the future to make them indifferent between the benchmark economy and the transition to the optimal steady state. As a consequence, the cost of the transition wipes out the welfare gains achieved in the steady state with the regressive tax policy.

5.2. Optimal Tax Reform along the Transition

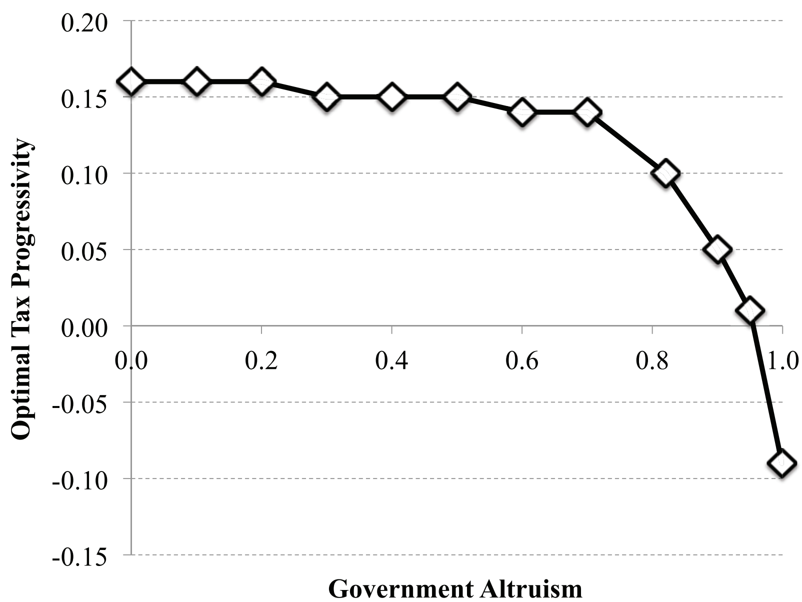
This raises the question which tax reform is optimal, starting in the U.S. status quo. The answer generally depends on how much the policy maker values future generations relative to the current generation. We begin with the case where the policy maker only cares to maximize the average welfare of agents that are alive at the time the tax reform is introduced ($\beta_g = 0$ in (6)). Recall that future generations are still valued in the policy maker's objective function due to parental altruism. In this case, the optimal policy is to slightly decrease progressivity, from 0.17 to 0.16. There are two underlying forces behind this result. On the one hand, the policy maker would like to increase long-run average output and consumption by implementing less progressive taxes. On the other hand, less progressive taxes imply lower consumption and leisure during the transition. They also lead to a less equitable distribution of income. Given the discount rate of the current generation, these forces balance each other. As a result, the optimal transition provides a welfare gain equivalent to 0.2% of consumption for the current generation. This is much lower than the gains suggested by the comparison of steady states only.

The utilitarian welfare function does not distinguish between ex-post inequality and ex-ante inequality. Therefore, it aims to equate marginal utility of consumption across agents not only to provide insurance for future variations in consumption, but also to eliminate pre-existing consumption differences. To better understand the insurance role of optimal taxation, we compute the optimal tax policy for the welfare function described in equation (7). The optimal progressivity in this case is $\tau = 0.10$. The reduction in the optimal progressivity highlights the egalitarian motive implicit in the utilitarian function. Since the elasticity of capital supply is relatively low in the short run, it can be taxed away from the wealthy and redistributed towards the income poor with little efficiency cost. This motive, akin to the intuition behind Chamley (1986), applies only if transition dynamics are considered when deciding the optimal tax policy.[19]

If, instead, the policy maker discounts the well-being of future generations less than the altruistic agents do, the long-run gains of regressive taxes realized by future generations become more important. We acknowledge that there is no "right" way to assign welfare weights across generations and report the results for various values of the policy maker's discount factor for future generations ($\beta_g \in [0, 1)$). The resulting progressivity of the optimal tax schedule for a utilitarian objective function is shown in Figure 4. As β_g increases, the optimal tax schedule becomes less progressive. For low values of β_g, the change in the optimal tax schedule is minor. This is due in part to the fact that each period, only a fraction μ of the population is newly born. As β_g approaches 1, the policy maker values future

[19] Since the steady-state level of capital is much more elastic, the two welfare functions yield similar results when only long-run equilibria are compared. The optimal progressivity is -0.08 when the welfare function in equation (7) is used.

Figure 4: Government Altruism and Optimal Tax Progressivity



Note.– Figure shows the progressivity of the optimal tax schedule (τ) as a function of the policy maker's discount factor applied to future generations ($\beta_g \in [0, 1)$).

generations as much as the current generation. In this extreme case, the long-run gains of reduced progressivity overtake the short-run costs along the transition. Consequently, the optimal tax progressivity approaches -0.09, which also maximizes steady-state welfare. If the policy maker values future generations as much as the altruistic parents do, ($\beta_g = \beta = 0.82$), the optimal tax system features $\tau = 0.10$, whereas a weight of $\beta_g = 0.96$ is needed for a flat tax system to be optimal.

The different levels of progressivity obtained under various values of β_g lead to substantially different distributions of the tax burden. Table 3 summarizes the steady-state distribution of the tax burden associated with each economy. In the optimal tax system obtained for $\beta_g = 0$, the average tax rate is 21.7% of taxable income.[20] The bottom 10% of the income distribution pays 7.1% of their disposable income in taxes while the top 10% pays 29.4%. This is similar to the existing tax system in the US. As the tax progressivity de-

[20]Taxable income is obtained by deducting the depreciation allowance from total income: $G/(Y - \delta K)$.

Table 3: The Distribution of Taxes and Income

Economy	Government Altruism (β_g)	Tax Policy (τ)	0 - 100	Income Percentiles <0.10	0.10–0.50	0.50–0.90	>0.90
				Average Tax Rates (%)			
Benchmark (U.S.)	–	0.17	21.8	6.7	15.0	23.1	29.7
Utilitarian	0.00	0.16	21.7	7.1	15.1	23.0	29.4
Altruistic	0.82	0.10	21.0	11.9	16.8	21.8	26.0
Flat Tax	0.95	0.00	20.0	20.0	20.0	20.0	20.0
Long-Run Optimum	1.00	-0.09	19.4	27.8	23.6	18.9	14.6
				Total Tax Share (%)			
Benchmark (U.S.)	–	0.17	100.0	1.2	18.2	50.8	29.9
Utilitarian	0.00	0.16	100.0	1.2	18.4	50.8	29.5
Altruistic	0.82	0.10	100.0	2.1	20.9	49.4	27.5
Flat Tax	0.95	0.00	100.0	3.7	25.7	48.0	22.6
Long-Run Optimum	1.00	-0.09	100.0	5.1	30.6	47.0	17.3

Note.– Table shows the steady-state distribution of the optimal tax rates and the resulting tax burden obtained under various welfare functions. The average tax rate in the model is determined by $1 - \lambda y^{-\tau}$ where y is pretax income and $\tau < 1$ denotes the progressivity of the tax system.

clines towards a flat tax system, the average tax rates incurred by different income groups converge. Note that lower progressivity leads to larger output and thus lower average tax rates in the aggregate. Recall that for the altruistic government that cares only about long-run outcomes, the optimal tax system is regressive with a τ of -0.09. The long-run equilibrium obtained under such a tax system features an average tax rate of 19.4. The bottom 10% pays 27.8% of their income in taxes, and the top 10% pays 14.6%.

The impact of declining progressivity on the poor is muted in terms of their contribution to the total government expenditure. For instance, while the average tax rate for the poor increases from 6.7% to 27.8%, the highest rate in the regressive system, the taxes collected from this group increase from 1.2% to 5.1% of total tax revenue. By contrast, the contribution of the top income group declines from nearly 30% to 17.3%.

To summarize, what constitutes an optimal tax policy depends crucially on household reactions to the policy, and on the relative welfare weights of different generations. The progressivity of taxes influences the need for self-insurance and capital accumulation. Regressive taxes are a tool that allows the policy maker to provide incentives for capital accumulation, with the added benefit of general equilibrium effects that favor the consumption-poor. Since the benefits of large policy changes require time to materialize, while costs are paid up front, the desirability of using this policy tool depends on how the policy maker weights the welfare of different generations. Hence, how regressive or progressive income taxes should be depends not only on current redistribution objectives, but crucially on the policy maker's welfare weights. If the policy maker discounts the well-being of future generations the way agents do, the well-being of generations that live early on dominates in the objective function. Consequently, it is optimal to sacrifice long-run steady-state welfare in exchange for current welfare improvements.

6. Discussion

The results highlight the role of transitional dynamics in the determination of an optimal redistribution system. When public policy is based solely on the comparison of steady states associated with different tax regimes, a utilitarian government may find it optimal to harness the effect of taxes on the demand for private insurance and, thereby, on equilibrium prices in order to improve social welfare. A regressive tax system achieves this by raising the wage rate and lowering the interest rate, giving the policy maker the option of eliminating the distortionary effects of a progressive system while keeping inequality under control. When the long-run outcomes of public policy are considered, this alternative is preferable to direct provision of social insurance via progressive taxation. The latter proves more costly as it also crowds out labor supply and savings.

When transitional dynamics are considered, direct redistribution through a progressive tax system is preferable instead. The cost of an immediate increase in inequality associated with less progressive taxes outweighs changes in long-run welfare. This is due to two reasons. First, reaching a steady state with a higher capital stock requires longer labor hours along the way. Second, since capital is inelastically supplied in the short run, the policymaker may find it optimal to raise the tax rate for high income groups, who rely more heavily on capital income. This is reminiscent of the result in Chamley (1986), where the government finds it optimal to initially confiscate the entire private capital stock before lowering the tax rate on capital to zero.

The sharp disparity between the optimal tax policy in these two scenarios illustrates the importance of the weighting of future generations relative to the current generation. If the policy maker only values future generations through parental altruism, a progressive tax system is optimal. If, in contrast, the policy maker assigns (almost) equal weights to all generations, welfare at the steady state outweighs short-run considerations. An alternative is to design intergenerational tax and transfer schemes to bring part of the increase in long-run average welfare to the current generation to compensate for longer hours. An implementation of such transfers may be possible if the policy maker can issue current debt against payments by future generations. How the possibility of debt issuance affects the optimal progressivity of income tax policy remains as a promising venue for future research.[21]

The optimal tax system may also differ if a reform is announced ahead of time, or if taxes can vary over time. In the former case, the policy maker would not be able to redistribute existing resources as effectively. Savings and thus the capital stock could decline in anticipation of a progressive reform. As a result, the policy maker would prefer a tax reform that is less progressive. Moreover, increased inequality is less costly if it only appears in the future, as dynasties gain time to prepare for the change in earnings risk. In addition, expected mean reversion of the fate of dynasties implies that attitudes to a reform will de-

[21] In recent work, Fehr and Kindermann (2014) and Krueger and Ludwig (2013) explore this avenue in OLG models.

pend more on its effect on average outcomes, and less on distributional effects, helping to increase the political support for a tax reform.[22] Finally, it seems likely that time-varying tax systems may make it easier to balance the concerns for inequality and welfare among currently alive generations with achieving long-run efficiency. A plausible conjecture is that a stream of tax policies with decreasing tax progressivity could help improve social welfare for a utilitarian government. An investigation of these issues is another promising direction for future research.

[22]An early contribution in this vein is Chamley (2001), who analyzes the long-run effects of pre-announced reforms to linear capital taxes in a related setting.

References

Acikgoz, O., 2013. Transitional dynamics and long-run optimal taxation under incomplete markets. Mimeo, Yeshiva University.

Aiyagari, R., 1994. Uninsured idiosyncratic risk and aggregate saving. Quarterly Journal of Economics 109, 659–684.

Aiyagari, S.R., 1995. Optimal capital income taxation with incomplete markets, borrowing constraints, and constant discounting. Journal of Political Economy 103, 1158–1175.

Barillas, F., Fernández-Villaverde, J., 2007. A generalization of the endogenous grid method. Journal of Economic Dynamics and Control 31, 2698–2712.

Bénabou, R., 2002. Tax and education policy in a heterogeneous agent economy: What levels of redistribution maximize growth and efficiency? Econometrica 70, 481–517.

Bewley, T., 1986. Contributions to Mathematical Economics: In honor of Gerard Debreu. New York: North Holland. chapter Stationary Monetary Equilibrium with a Continuum of Independently Fluctuating Consumers. pp. 79 – 102.

Black, S., Devereux, P.J., 2011. Handbook of Labor Economics. Elsevier Science B. V.. volume 4B. chapter Recent Developments in Intergenerational Mobility. pp. 1487–1541.

Blundell, R., MaCurdy, T.E., 1999. Labor Supply: A Review of Alternative Approaches. North Holland, Amsterdam. chapter 27. Handbook of Labor Economics, pp. 1559 – 1694.

Blundell, R., Pistaferri, L., Saporta-Eksten, I., 2012. Consumption inequality and family labor supply. NBER Working Paper 18445.

Cagetti, M., De Nardi, M., 2006. Entrepreneurship, Frictions, and Wealth. Journal of Political Economy 114, 835–870.

Carroll, C.D., 2006. The method of endogenous gridpoints for solving dynamic stochastic optimization problems. Economics Letters 91, 312–320.

Castaneda, A., Díaz-Giménez, J., Ríos-Rull, J.V., 2003. Accounting for the us earnings and wealth inequality. Journal of Political Economy 111, 818–857.

Chamley, C., 1986. Optimal taxation of capital income in general equilibrium with infinite lives. Econometrica 54, 607–22.

Chamley, C., 2001. Capital income taxation, wealth distribution and borrowing constraints. Journal of Public Economics 79, 55–69.

Cho, J., Cooley, T.F., 1994. Employment and hours over the business cycle. Journal of Economic Dynamics and Control 18, 411 – 432.

Conesa, J.C., Kitao, S., Krueger, D., 2009. Taxing Capital? Not a Bad Idea After All! The American Economic Review 99, 25–48.

Conesa, J.C., Krueger, D., 2006. On the optimal progressivity of the income tax code. Journal of Monetary Economics 53, 1425–1450.

Cutler, D., Gruber, J., 1996. Does public insurance crowd out private insurance? The Quarterly Journal of Economics 111, 391.

Davila, J., Hong, J.H., Krusell, P., Ríos-Rull, V., 2012. Constrained efficiency in the neoclassical growth model with uninsurable idiosyncratic shocks. Econometrica 80, 2431–2467.

Diaz-Gimenez, J.A., Glover, A., Ríos-Rull, V., 2011. Facts on the distributions of earnings, income, and wealth in the United States: 2007 update. Federal Reserve Bank of Minneapolis Quarterly Review 34, 2–31.

Domeij, D., Heathcote, J., 2004. On the distributional effects of reducing capital taxes. International Economic Review 45, 523 – 554.

Erosa, A., Koreshkova, T., 2007. Progressive taxation in a dynastic model of human capital. Journal of Monetary Economics 54, 667–685.

Farhi, E., Werning, I., 2007. Inequality and social discounting. Journal of Political Economy 115, 365–402.

Farhi, E., Werning, I., 2012. Capital taxation: Quantitative explorations of the inverse euler equation. Journal of Political Economy 120, 398–445.

Feenberg, D.R., Coutts, E., 1993. An introduction to the taxsim model. Journal of Policy Analysis and Management 12, 189–194.

Fehr, H., Kindermann, F., 2014. Optimal taxation with current and future cohorts. CESifo Working Paper.

Golosov, M., Kocherlakota, N., Tsyvinski, A., 2003. Optimal indirect and capital taxation. Review of Economic Studies 70, 569–87.

Golosov, M., Tsyvinski, A., 2007. Optimal taxation with endogenous insurance markets. Quarterly Journal of Economics 122, 487–534.

Guner, N., Kaygusuz, R., Ventura, G., 2014. Income taxation of us households: Facts and parametric estimates. Review of Economic Dynamics in press.

Heathcote, J., Perri, F., Violante, G.L., 2010. Unequal we stand: An empirical analysis of economic inequality in the united states, 1967–2006. Review of Economic Dynamics 13, 15–51.

Heathcote, J., Storesletten, K., Violante, G., 2014. Optimal tax progressivity: An analytical framework. Federal Reserve Bank of Minneapolis Research Department Staff Report 496.

Hubbard, R.G., Skinner, J., Zeldes, S.P., 1995. Precautionary Saving and Social Insurance. Journal of Political Economy 103, 360–399.

Huggett, M., 1993. The risk-free rate in heterogeneous-agent incomplete-insurance economies. Journal of Economic Dynamics and Control 17, 953–69.

Kocherlakota, N., 2005. Zero expected wealth taxes: A mirrlees approach to dynamic optimal taxation. Econometrica 73, 1587 – 1621.

Krebs, T., 2007. Job displacement risk and the cost of business cycles. American Economic Review 97, 664 – 686.

Krueger, D., Ludwig, A., 2013. Optimal Progressive Taxation and Education Subsidies in a Model of Endogenous Human Capital Formation. Penn Institute for Economic Research Working Paper.

Krueger, D., Perri, F., 2011. Public versus private risk sharing. Journal of Economic Theory 146, 920–956.

Lucas, R.E.J., 1987. Models of Business Cycles. New York: Blackwell.

Marcet, A., Obiols-Homs, F., Weil, P., 2007. Incomplete markets, labor supply and capital accumulation. Journal of Monetary Economics 54, 2621–2635.

McGrattan, E.R., Prescott, E.C., 2010. Unmeasured investment and the puzzling us boom in the 1990s. American Economic Journal: Macroeconomics 2, 88–123.

Mirrlees, J.A., 1971. An exploration in the theory of optimum inome taxation. Review of Economic Studies 38, 175–208.

Mulligan, C., 1997. Parental Priorities and Economic Inequality. University of Chicago Press, Chicago, IL.

Piketty, T., Saez, E., 2007. How progressive is the U.S. federal tax system? a historical and international perspective. Journal of Economic Perspectives 21, 3–24.

Prescott, E.J., 2004. Why do americans work so much more than europeans? Federal Reserve Bank of Minneapolis Quarterly Review 28, 2–13.

Ramsey, F., 1927. A contribution to the theory of taxation. Economic Journal 37, 47 – 61.

Ríos-Rull, J., Attanasio, O., 2000. Consumption smoothing in island economies: Can public insurance reduce welfare? European Economic Review 44, 1225–1258.

Seshadri, A., Yuki, K., 2004. Equity and efficiency effects of redistributive policies. Journal of Monetary Economics 51, 1415–1447.

Solon, G., 1992. Intergenerational income mobility in the united states. American Economic Review 82, 393–408.

Solon, G., 1999. Handbook of Labor Economics. Elsevier Science B. V.. volume 3. chapter 29.

Storesletten, K., Telmer, C.I., Yaron, A., 2004. Consumption and risk sharing over the life cycle. Journal of Monetary Economics 51, 609–633.

Appendix: Computational Algorithm

Steady State. 1. Choose a grid for asset holdings. 2. Guess a value for λ_{SS}. 3. Guess a value for r_{SS}. 4. Compute the implied capital-labor ratio and the wage rate. 5. Solve the household's problem for saving and labor supply. To do so, we iterate until convergence on an alternation of the endogenous grid point method for the saving policy (Carroll, 2006; Barillas and Fernández-Villaverde, 2007) and a bisection on the first order condition for labor supply. On some rare occasions where this method does not converge, we use value function iteration for one step. 6. Compute the stationary distribution of assets and z implied by the law of motion of z and the savings policy function. 7. Using this distribution and the household policies, compute the aggregate capital stock and labor supply. 8. Using these two terms, compute the net marginal product of capital. This is the *implied* (by policies) steady state interest rate. 9. Check whether the implied r equals the guess. If not, compute a new guess for r as a convex combination of the last guess and the implied value, and return to step 3. 10. Compute the *implied* λ using the government budget constraint and household incomes implied by their policies. Check whether the implied λ equals the guess. If not, compute a new guess for λ as a convex combination of the last guess and the implied value, and return to step 2.

Transition. 1. Compute the initial and final steady states. 2. Choose a length T for the transition. 3. Guess a path for $\{r_t, \lambda_t\}_{t=1}^{T}$. 4. Compute implied capital-labor ratios and wage rates for each t. 5. Using V_{SS2}, solve for the optimal labor and saving policies at each t, given prices and policies, moving backwards from $t = T$ until $t = 1$. 6. Because capital is predetermined, the distribution of capital in period 1 is already known. Use it to compute K_1. Use the joint distribution of capital and z in period 1 to compute aggregate labor supply in period 1. Using these two terms, compute the net marginal product of capital in period 1. This is the *implied* (by policies) interest rate for period 1. 7. Compute income and disposable income in period 1, and obtain *implied* λ_1 from the government budget constraint. 8. Using the savings policy for period 1 obtained in step 5 and the distribution of capital in period 1, compute the distribution of capital in period 2. 9. Repeat steps 6 to 8 for all periods until period T. 10. Check whether implied r_t and λ_t, $t = 1...T$, equal the guesses. If yes, the problem is solved. If not, compute new guesses for r_t and λ_t for each t as a convex combination of the last guess and the implied value, and return to step 4. 11. Verify that the solution is not sensitive to T.

Note. The transition algorithm uses the same grid for capital as the steady state algorithm.