Supply Side Interventions and Redistribution*

Teresa Garcia-Milà
Albert Marcet
Eva Ventura†

1st version July 1995
July 2008

Abstract

We evaluate the effect on welfare of shifting the burden of capital income taxes to labor taxes in a dynamic equilibrium model with heterogeneous agents and constant tax rates. We calibrate and simulate the economy; we find that lowering capital taxes has two effects: *i)* it increases efficiency in terms of aggregate production, and *ii)* it redistributes wealth in favor of those agents with a low wage/wealth ratio. When the parameters of the model are calibrated to match the distribution of income in terms of the wage/wealth ratio, the redistributive effect dominates, and agents with a high wage/wealth ratio would experience a large loss in utility if capital income taxes were eliminated.

*We are thankful to J. Geweke, E. McGrattan, A. Braun, E. Prescott, L. Christiano, O. Attanasio, M. Reiter, N. Roubini, D. Epple and C. Sims for helpful conversations. Katharina Greulich has provided invaluable research assistance. Part of Marcet's work was done while he was a visitor at the Federal Reserve Bank of Minneapolis. Support from the DGICYT, CIRIT, Instituto de Estudios Fiscales, CREI and the Barcelona Economics' program of CREA is greatly appreciated. All errors are our own.

†Garcia-Milà and Ventura are at Universitat Pompeu Fabra. Marcet is at Institut d'Anàlisi Econòmica CSIC and CEPR.

1 Introduction

A large part of the literature on dynamic taxation in equilibrium models with rational expectations has reached the conclusion that capital taxes should be abolished or, at the very least, severely reduced. Chamley (1986) showed that in a dynamic equilibrium model with proportional taxes, full commitment and time-varying taxes, it was optimal to suppress capital taxes in the long run. This reduction in capital taxes would promote aggregate investment, increase production and consumption in the long run.

This result has been shown to be robust to many extensions.[1] In particular, it is robust to the introduction of heterogeneity: even if agents are heterogeneous optimal policies drive capital taxes to zero in the long run.[2] In this way the study of capital taxation in dynamic rational expectations models has provided rigorous ground for an old idea in economics: a decrease in capital taxes would increase the size of the pie and, perhaps, make everybody better off.

The reduction of capital taxes is not a purely academic issue, it has been at the forefront of policy discussions. Some countries have recently reduced capital gains taxes or corporate taxes To mention a few, Spain, France, Sweden and the US . The recent economic success of Ireland is often linked to lower capital taxes. Indeed, most measures show that capital taxes were extremely high, but that they have been going down in the last two decades. Carey and Tchilinguirian (2000), with estimates for the OECD countries for the period 1980-97, conclude that there has been a shift in the relative tax burden from capital to labor, with an average annual decrease of -.2% in capital taxes, and an increase of .3% on labor taxes. For the US these rates are -.5% and .2% respectively.

Chamley's result is only about *long run* tax rates: it is well known that optimal capital taxes are not zero in the transition to the steady state. As shown in Jones, Manuelli and Rossi (1993) the transition of optimal taxes shows very large oscillations through time. Optimal taxes can take extreme values in different periods, and the exact shape of the transition is highly dependent on the exact model at hand, making it difficult to implement the

[1] For a review of the extensions see the relevant chapters of Ljunqvist and Sargent (2004) and Chari and Kehoe (1999).

[2] The result is obtained in Chamley (1986) and Judd (1985) and (1987). A proof for the model considered in the current paper where no lump sum transfers ara available is found in Atkeson, Chari and Kehoe (1999).

correct Ramsey tax policy in the real world. Therefore it is of interest to study the effect of implementing policies with simpler dynamics, in particular, policies with constant tax rates. Inspired by the long run results of Chamley, one could consider the effect of abolishing capital taxes and to set labor taxes to a new constant level, high enough to keep the same level of government spending. Lucas (1990) performed exactly this experiment in a neoclassical dynamic model of capital accumulation and he found that abolishing capital taxes and shifting the burden of tax revenue to labor taxes was welfare improving. Cooley and Hansen (1992) confirmed these results even when considering inflation tax and consumption taxes.

Lucas (1990), and Cooley and Hansen (1992) used a model with homogeneous agents. Therefore, they could not address issues of equity and redistribution that immediately come to mind when the issue of capital vs labor taxes is brought to discussion. Given the highly skewed distribution of wealth observed in the real world we would expect large redistributive effects of abolishing capital taxes. The object of the current paper is to study the effects of abolishing capital taxes in a model with heterogeneous agents. In this way we can address both issues of efficiency and equity.

We keep the model as close as possible to that of Chamley. Therefore we consider a model of capital accumulation, infinitely-lived agents, flexible prices, proportional capital and labor taxes, complete markets and competitive equilibrium. We rule out redistributive lump sum taxes, as these would render the redistributive issue irrelevant and such taxes are impossible to implement in the real world. We also consider agents that can both save and work, as in the data the vast majority of agents (excluding retired) do so. We calibrate our model to observed heterogeneity of agents in a relevant way for the exercise at hand. We find the usual result that a reduction in capital taxes enhances economic activity: wages, aggregate investment, aggregate consumption and aggregate output all increase by a significant amount. Nevertheless, abolishing capital taxes also changes the distribution of wealth since it increases the disposable income of capital-rich agents in a major way; the redistributive effect is so important that the utility of agents with a high wage/wealth ratio decreases dramatically; only consumers with a low wage/wealth are better off. The effects on individual welfare are very large: the lowest quintile of the population would suffer a loss of between 20% and 60% (depending on the calibration). Furthermore, depending on the calibration, either 40% or 60% of the population would loose from the reform. This means that, with heterogeneous agents, the transitional path of the optimal (time-varying) tax policy is crucial in reaching the optimal Chamley/Judd allocations where all agents improve, and that simply abolishing capital taxes has very large redistributive effects.

Some papers have shown how it may be difficult to implement Ramsey policies due to time inconsistency. For example, Klein, Krusell and Ríos-Rull (2007) show how a time consistent policy would involve capital taxes that are quite high in the long run. One possible conclusion from these observations is that issues such as lowering capital taxes should be written in the constitution. Our results would say the new constitution would have to be written very carefully in order to be approved. The median voter is likely to disagree with a change in the constitution stating that capital taxes are lowered in one step and that they can never be changed afterwards.

Since we are extrapolating the behavior of the economy into an area where no observations are available, the answer to this issue is both highly dependent on theoretical and empirical elements introduced in the model. In the paper we provide a careful discussion of two aspects that are of particular importance for the exercise at hand, namely the relevant aspects of the wealth distribution and the elasticity of labor. In the empirical literature on inequality it is standard to focus on either the distribution of wealth or the distribution of income. We argue that, in fact, the whole *joint* distribution of wealth, capital and labor income across the population matters. In particular, we show that for our policy analysis the relevant dimension of this distribution is the dispersion of the *wage/wealth ratio* across agents. Therefore we calibrate the wage and wealth of our agents to the observed distribution of the wage/wealth ratio. Another key aspect in the calibration are the parameter values and functional forms that concern the elasticity of labor, since this will influence the efficiency cost of the higher labor taxes that are needed to compensate for the lost capital tax revenue. We argue that the standard neoclassical model does not allow to match both the variability of hours worked across time and across agents. Since we are particularly concerned about agents' heterogeneity we choose a highly inelastic labor supply to roughly match the cross-section observations. Therefore a shortcoming of our model is that it predicts a time series variability of aggregate hours worked much too low. Finding a model that matches both cross section and time series variability of hours worked would be ideal but it is a puzzle that we do not address in this paper.

We use a standard neoclassical model of capital accumulation in order to stay as close as possible to the original setup of Chamley.[3] This allows us to show in isolation the effects of introducing heterogeneity in the setup of Lucas (1990), and Cooley and Hansen (1992), and in a model where the Chamley-Judd result holds. Other papers have analyzed related issues.[4] These papers confirm that our main finding is robust to many extensions: suppressing capital taxes has large redistributive effects that would strongly decrease the welfare of large parts of the population. To our knowledge other papers available do not incorporate the calibration of the distribution of the wage/wealth ratio that we have emphasized.

Along the way, we reexamine the result of Chari Christiano and Kehoe (1994) that suppressing capital taxes in a model with a representative agent and high risk aversion would be undesirable. We find that if the calibration maintains an empirically plausible capital output ratio, a homogeneous agent would still prefer suppressing capital taxes. However, in a model with heterogeneous agents and high risk aversion, the redistributive effects of suppressing capital taxes are very high.

The layout of the paper is as follows. The model is presented in section 2. Section 3 discusses issues pertaining to parameter calibration using data from the US economy. Section 4 presents the results derived from the simulations. Section 5 performs sensitivity analysis to parameter values and to the introduction of uncertainty. The conclusion ends the main paper. The appendices discuss the details of the calibration using PSID data set, how to convert the model with growth into one in deviations from trend, the numerical algorithms used and the equilibrium conditions.

---

[3] This paper is a revised version of our 1995 working paper. The current version has five agents, the main analysis is performed with a deterministic model, there are more robustness checks of the results to parameter changes, and we have added the analysis for the high risk aversion case.

[4] Correia (1999) shows analytically the source of redistributive effects in a model with aggregation, Domeij and Heathcote (2004) use a model with incomplete markets and focus on the effects of idiosincratic uncertainty. Flodén (2007) studies a model where the transition of capital taxes is found optimally for one of the agents in the model. Maliar and Maliar (2001) derive aggregation results, calibrate the model with 8 heterogeneous groups of agents and compare the results to those in García-Milà et al. (1995). Some work has been done on models of overlapping generations, for example Altig and Carlstrom (1999) consider taxation and inequality in an equilibrium model with overlapping generations, although different from our work they include redistributive transfers and do not distinguish between capital and labor taxes. Closer to our work Conesa, Kitao and Krueger (2007) show that in a model with overlapping generations and idiosyncratic uninsurable risk it can be ex-ante optimal to have high capital taxes.

2 The Model

In this section we describe a simple neoclassical growth model with heterogeneous agents, endogenous production, labor choice, exogenous growth,[5] and government spending. Government can only use distortionary capital and labor taxes. Agents differ both in terms of their human, and non-human wealth. The benchmark model is deterministic. We introduce aggregate uncertainty in section 5 in order to address some empirical and robustness issues.

2.1 Consumer, Firm, and Government Behavior

Assume that n infinitely-lived consumer types indexed by $j = 1, 2, \ldots, n$ derive utility from consumption and leisure, and they are endowed with one unit of time every period. The number of each type of agents is normalized to $1/n$. They receive income from working and from renting their capital. All agents can work and accumulate (or divest) capital. Agents are heterogeneous in their endowment of labor productivity and initial capital stock, and their labor and capital incomes are taxed at constant rates τ^l and τ^k.

Consumers of type j solve the following maximization problem:

$$\max_{\{x_{jt}\}} \sum_{t=0}^{\infty} \delta^t \left[u(c_{j,t}) + v(l_{j,t}, \mu^t) \right]$$

$$\text{s.t.} \qquad c_{j,t} + k_{j,t} - k_{j,t-1} = \phi_j \ \mu^t \ w_t \ l_{j,t}(1-\tau^l) + k_{j,t-1}(r_t - d)(1-\tau^k) \tag{1}$$

$$k_{j,-1} \text{ given}$$

where $\{x_{j,t}\} \equiv \{c_{j,t}, l_{j,t}, k_{j,t}\}_{t=0}^{\infty}$ are the choice variables of the consumer.

We assume separability in time and in the consumption-leisure decision. Here, $c_{j,t},\ k_{j,t},\ l_{j,t}$ denote consumption, capital stock and hours worked of agent j at time t; $w_t,\ r_t$ denote prices of efficiency units of work and capital rental, normalized in terms of the consumption good of the period. Agent j produces ϕ_j efficiency units per hour worked. Labor productivity is assumed to grow exogenously at the rate μ; the wage obtained per hour worked in

[5] Introducing growth explicitely is important in order to quantify the efect of depreciation allowances. This is because in the stationary version of the model total investment is no longer equal to gross investment, therefore the size of the tax base is not the same as if the basis of the analysis would be the stationary version of the model.

period t by agent j is, therefore, $\phi_j \ \mu^t \ w_t$.[6] Since we concentrate our study on issues of distribution, our agents only differ in their initial wealth $k_{j,-1}$ and their efficiency of labor $\phi_j > 0$; these are normalized so that $\frac{1}{n}\sum_{j=1}^{n}\phi_j = 1$. Parameters δ, d are in the interval [0,1], they stand for the discount factor of future utility and the depreciation rate of capital. Notice that taxes on capital are net of depreciation allowances. Functions u and v are differentiable and satisfy the appropriate Inada conditions to insure interior solutions; $u(\cdot)$ and $v(\cdot, \mu)$ are strictly concave; $u(\cdot)$ and $v(l, \cdot)$ are strictly increasing and $v(\cdot, \mu)$ is strictly decreasing. Notice that capital holdings could be negative if the agent decides to go in debt.[7]

There is one representative firm that maximizes period-by-period profits; it manages a production technology, rents capital at a price r_t and hires efficiency units of labor at a wage w_t to solve

$$\begin{array}{lrcl} \max\limits_{(y_t, e_t, k_{t-1})} & & & y_t - w_t e_t - r_t k_{t-1} \\ \text{s.t.} & y_t & = & F(k_{t-1}, \ e_t) \end{array} \tag{2}$$

where y_t represents output, k_{t-1} the demand of capital, and e_t the demand for efficiency units of labor. F is the production function gross of depreciation, strictly concave and homogeneous of degree one.

Since total efficiency units of labor supplied is $\frac{1}{n}\left(\sum_{j=1}^{n} \phi_j l_{j,t}\right)\mu^t$ all variables grow at a rate μ in the steady state except labor, which is constant in steady state. The fact that the size of each group of agents is normalized to $1/n$ and the normalization $\frac{1}{n}\sum_{j=1}^{n}\phi_j = 1$ guarantees that by setting $\phi_i = \phi_j$ and $k_{i,-1} = k_{j,-1}$ for all $i, j = 1, 2, \ldots, n$ we are back to the homogeneous agent model in Lucas (1990).

We now discuss the constraints of government fiscal policy. Government spending is exogenous but grows at the same rate as output, so the sequence

[6] Introducing the trend of labor productivity (μ^t) in the utility function is a standard way to insure a non-degenerate solution for hours worked in the long-run in the presence of growth. This formulation has been controversial. Some economists have argued that this is artificial, while others have argued that it is consistent with assuming that higher human capital yields higher utility from leisure. This controversy is not relevant for our benchmark calibration with log utility, where the term μ^t actually is absent from the utility function. We only need the term μ^t in the utility function for some of the robustness exercises in section 5.

[7] As usual, some additional lower bound on (possibly negative) capital holding has to be introduced in order to rule out Ponzi schemes. The same will be true for the budget constraint of the government.

of government consumption is given by $g_t \equiv \mu^t g$ for a given constant g.[8] Tax revenues accrue from constant capital and labor tax rates τ^k, τ^l. Government can save or dissave by borrowing or lending at equilibrium interest rates. It is well known this is equivalent with assuming that the government has (possibly negative) capital stock holdings k_t^g. This amounts to the following budget constraint at period-t

$$g_t + k_t^g - k_{t-1}^g = (r_t - d)\left(\frac{1}{n}\sum_{j=1}^{n} k_{j,t-1}\right)\tau^k + w_t\, e_t \tau^l + (r_t - d)(1-\tau^k)\, k_{t-1}^g \quad (3)$$

Initial government savings k_{-1}^g are given.

2.2 Equilibrium

We assume competitive equilibrium. As usual, an equilibrium is a sequence for prices and allocations, and a government policy (g, τ^k, τ^l), such that when consumers maximize utility and firms maximize profits taking prices and government policy as given, they choose equilibrium allocations that clear all markets and the budget constraint of the government is satisfied.

The equations determining equilibrium are as follows. Market clearing in capital, labor and consumption good are given, for all t, by

$$k_t^g + \frac{1}{n}\sum_{j=1}^{n} k_{j,t} = k_t \quad (4)$$

$$\frac{1}{n}\left(\sum_{j=1}^{n} \phi_j l_{j,t}\right)\mu^t = e_t \quad (5)$$

$$\frac{1}{n}\sum_{j=1}^{n} c_{j,t} + g_t + k_t - (1-d)k_{t-1} = y_t \quad (6)$$

With interior solutions, the first order conditions for capital, and labor

[8] Since we maintain g constant across policy experiments, the equilibrium computed and the welfare gains discussed in sections 4 and 5 are consistent with a model where government spending enters the utility function or the production function. To keep notation simple, we write the paper as if government spending has no productive use.

choice in the consumer's problem are

$$u'(c_{j,t}) = \delta\, u'(c_{j,t+1}) \left(1 + (r_{t+1} - d)(1 - \tau^k)\right) \tag{7}$$

$$-\frac{v'(l_{j,t}, \mu^t)}{u'(c_{j,t})} = w_t\ (1 - \tau^l)\ \mu^t \phi_j \tag{8}$$

for all t and j. Here, $v' \equiv \frac{\partial v}{\partial l}$. These are the familiar conditions setting the intertemporal marginal rate of substitution of consumption (between leisure and consumption) equal to the price of capital (labor) net of taxes.

As usual, equilibrium factor prices equal marginal product to set $r_t = F_k(k_{t-1},\ e_t)$ and $w_t = F_e(k_{t-1},\ e_t)$.

Equation (7) implies that for some constants λ_j

$$\frac{u'(c_{n,t})}{u'(c_{j,t})} = \frac{\phi_j\ v'(l_{n,t}, \mu^t)}{\phi_n\ v'(l_{j,t}, \mu^t)} = \lambda_j \qquad \text{for all } \ t, \ \text{all} \ \ j = 1, \ldots, n-1 \tag{9}$$

For constant relative risk aversion (CRRA) utility of consumption this is the familiar condition that under complete markets and common discount factors the share of consumption is constant through time.

Finally, substituting (7) and (8) we obtain the present value formulation of the consumers' budget constraints

$$\sum_{t=0}^{\infty} \delta^t \frac{u'(c_{n,t})}{u'(c_{n,0})} \left(c_{j,t} - w_t\ \mu^t\ \phi_j\ l_{j,t}(1 - \tau^l)\right) = \tag{10}$$
$$k_{j,-1}(1 + (r_0 - d)(1 - \tau^k)) \quad \text{for} \ \ j = 1, 2, \ldots, n$$

The budget constraint of the government is guaranteed by Walras' law and, therefore, can be ignored. This equation guarantees that the individual capital holdings can be ignored.

Equation (9) means that given $\{c_{n,t}, l_{n,t}\}$ and the λ's we can find consumption and labor for all agents $j = 1, ..., n-1$. Therefore, for each feasible policy choice $\left(\tau^l, \tau^k, g\right)$ all equilibrium quantities can be obtained by finding $\{c_{n,t}, l_{n,t}, k_t\}_{t=0}^{\infty}$ and λ's that satisfy (6), (7) for $j = n$, equation (8) for $j = n$, and the present value budget constraints (10) for $j = 1, ...n$. This reduces the number of variables and equations that need to be found to compute an equilibrium.[9,10]

[9] More precisely: equations (7) and (8) for $j = 2, \ldots, n$, period-t budget constraints (1) and (3) and the budget constraint of the government can be ignored.

[10] Appendix 4 gives a formal proof of this simplification for the stochastic case and the results, therefore, also apply to the non-stochastic case.

This concludes the description of all equilibrium conditions.

3 Calibration, some stylized facts and some analytic results

For our calibration we assume the following functional form of the utility function:[11]

$$u(c) = \frac{c^{\gamma_c+1}}{\gamma_c+1} \quad \text{and} \quad v(l,\mu^t) = B\,\frac{(1-l)^{\gamma_l+1}}{\gamma_l+1}\mu^{t(\gamma_c+1)} \tag{11}$$

for $\gamma_c, \gamma_l < 0$ and $B > 0$, and we assume that hours worked satisfy $0 \leq l_{j,t} \leq 1$. Notice that, since we will choose $\gamma_c = -1$ in the benchmark calibration the term μ disappears from the utility function in that case.

As usual we use a Cobb-Douglas production function $F(k_{t-1},\ e_t) = \mu^{\alpha}\ A\ k_{t-1}^{\alpha}\ e_t^{1-\alpha}$.

The effects of a tax reform are highly dependent on parameter values. Therefore, we need to use parameter values that can arguably represent the behavior of actual economies in a dimension that is relevant for our exercise. In the rest of this section we describe the criteria that guided our choice of parameter values in the benchmark economy. These parameters are summarized in Table 1.

3.1 Preference, technology and policy parameters

To insure comparability with the rest of the literature and to match various empirical regularities that are successfully explained by neoclassical growth models many parameters are chosen in a standard way.

[11] Correia (1999) and Flodén (2007) obtain interesting results using Greenwood Herkovicz and Huffman (GHH) preferences. These preferences guarantee aggregation in models with capital and labor taxes. The analysis is simplified, since aggregate quantities can be obtained without reference to the heterogeneity parameters. However, even if aggegation obtains, one has to resort to numerical simulations to find the welfare effects of the tax reform that we consider, so we would not save much numerical work by resorting to these preferences. Furthermore, it is well known that standard GHH preferences imply that status quo hours worked are zero in our model with growth. Finally, GHH preferences have the counterfactual implication that the wealth elasticity of hours worked is zero. This elasticity plays an important role in the response of hours worked to the elimination of capital taxes. For these reasons we use more standard preferences where aggregation does not obtain for a reasonable calibration of the heterogeneity parameters.

We choose log utility, $\gamma_c = -1$. This represents a low level of risk aversion but it is the value most commonly found in studies of fiscal policy. In this case we see from (9) that λ_j gives exactly the consumption ratio relative to agent n :

$$\frac{c_{j,t}}{c_{n,t}} = \lambda_j \qquad j = 1, \ldots, n-1 \quad . \tag{12}$$

As usual, B is chosen so that the representative agent works 1/3 of his time endowment in the steady state corresponding to the tax rates previous to the policy reform. Also, α is chosen to match the labor share of income. Depreciation rate, discount rate of utility, and growth rate are set to the usual values for quarterly data.

The policy parameters (τ^l, τ^k, g) are chosen to match measured average *effective* marginal tax rates. There is a long literature on this measurement. Papers vary in the method employed to measure these taxes, in the sample used, in the introduction of depreciation allowances and growth. We use McGrattan, Rogerson and Wright (1997) estimates of τ^k =.57 and τ^l =.23 for the period 1947-87, who follow the procedure of Joines (1981). These values are not too different from the ones estimated for the US in Carey and Tchilinguirian (2000), who updating the Mendoza, Razin and Tesar (1994) methodology obtain estimates of around .5 for capital tax and .22 for labor tax for the period 1980-97.[12] In any case, we will discuss in detail the sensitivity of our results to the value of τ^k.

Government spending is selected to balance the government budget constraint in status quo at steady state before the reform.[13]

Initial capital is set at the steady state of the status quo policy, that is,

[12] The rate of $\tau_k = .57$ is not as high as it may appear, since it is applied to income after depreciation allowances and since this is the sum of all taxes on capital income paid by consumers and firms. In any case, there is considerable disagreement on the relevant level of labor and income taxes, specially on the level of the capital tax. Feldstein, Dicks-Mireaux and Poterba [1983] obtain estimates of τ^k that range between .55 and .85 for the period 1953-1979. Cooley and Hansen use a lower tax rate, setting $\tau^k = .5$, (this number is based on Joines [1981] with the data ending in 1979), and they do not substract growth from the depreciation allowances; Chari, Christiano and Kehoe use $\tau^k = .27$; Lucas [1990] considers capital and labor taxes of .4; Greenwood, Rogerson and Wright [1995] set $\tau^k = .70$.

[13] Since we are interested in the effects of substituting capital taxes by labor taxes, and in keeping with the practice in Lucas (1990) and Cooley and Hansen (1992), we will only consider government spending that is financed from these two taxes. Therefore, total government spending in our model will be lower than the one actually observed.

assuming that the previous policy has been in place for a very long time. We choose $A = 1$ in the production function, as is standard in the literature, even though this will mean that the capital/output ratio will be lower than usual.[14]

Initial government debt is set to $k^g_{-1} = -2$. Since output is close to 1 and the model is calibrated to quarters this amounts to choosing a yearly debt/output ratio of about fifty percent in the status quo.

3.2 Heterogeneity parameters

The parameters that determine agents' heterogeneity, namely the productivity of labor ϕ_j and initial levels of wealth $k_{j.-1}$, are key to the outcome of the policy reform we consider. Therefore it is important to calibrate these parameters so as to capture appropriately the actual joint distribution of wage and wealth across agents. We focus on those aspects of this distribution that are key for the policy outcome.

We argue that the relevant dimension to be matched is the distribution of wage/wealth ratios across agents. This is because two agents with the same wage/wealth ratio are likely to both loose or gain from the reform we consider, even if one of them has a much higher total income than the other. To demonstrate this point we show a concrete example. Consider the case where the wage/wealth ratio is constant across all agents:

$$\frac{\phi_i}{k_{i,-1}} = \frac{\phi_j}{k_{j,-1}} \quad \text{for all} \quad i, j = 1, 2, \ldots, n \tag{13}$$

That is, an agent who is twice as productive is also twice as wealthy. Also, for simplicity, consider $\mu = 1$ and $k^g_{-1} = 0$.

It can be easily checked that for any set of tax rates equilibrium allocations in this example satisfy

$$\frac{c_{i,t}}{c_{j,t}} = \frac{\phi_i}{\phi_j} \qquad , \qquad l_{i,t} = l_{j,t} \qquad \text{for all} \quad t, i, j.$$

[14]Table 4 shows that the capital/output ratio of the status quo economy is below seven, lower than the usual values of ten or twelve for a quarterly model. This lower capital/output ratio is due to the large capital taxes combined with the standard $A = 1$. Changing A so as to match the capital/output ratio does not change the results significantly.

In words, all agents work the same but an agent twice as productive (and, under (13), twice as wealthy) consumes and saves twice as much each period.

It is clear that, in this case, the ratio λ_j is equal to $\frac{\phi_j}{\phi n}$, therefore this ratio is independent of tax rates. It follows that any gain or loss from alternative tax policies affects equally the profile of consumption and leisure of all agents. If agent i consumes double than agent j before the reform, agent i will continue to consume double than j after the reform.

If (13) was a good approximation to the actual distribution of wealth and wages all agents would experience a similar gain from the tax reform we consider. In this case introducing heterogeneity in the model provides no new insights. On the other hand, if we find a lot of dispersion of wage/wealth ratios in actual data some agents may gain and others may loose from suppressing capital taxes. Therefore, we should examine if (13) is a good approximation to the empirical distribution of income.

For this purpose we examine the joint distribution of wealth and wages in actual data. Figure 1 plots wages against wealth for different households computed from the PSID.[15] Each dot represents the wage and wealth of a family in the sample. If (13) was a good approximation to actual data most dots would be located near a straight line going through the origin (a "ray"). It is obvious, however, that the actual distribution is not grouped along one ray. The dispersion of wage/wealth ratios is very high and therefore abolishing capital taxes may affect different agents differently.

The issue is, then, how to introduce the relevant aspects of the distribution of wages and wealth in the model in a parsimonious way. Agents located either in the upper left corner or in the lower right corner of this figure are both "rich", but those agents in the upper left corner are likely to loose from the abolition of capital taxes because most of their income comes from labor, which will be taxed more heavily after the reform. Agents with a similar wage/wealth ratio either all gain or all loose, regardless of their total wealth.

To give some names to the situation: it is not that important, for the analysis in this paper, to distinguish between a yuppie and a low-qualified worker. They might have a very different level of income but both have a high wage/wealth ratio. It is important, however, to distinguish between the yuppie and a large land-owner who has zero labor income: they both have a high total income but they have very different wage/wealth ratios. The large land-owner is likely to gain from the reform we consider while the yuppie is

[15] The details on how this Figure has been constructed are in Appendix 1.

likely to loose.

In most studies of the wealth distribution the usual criterion is to classify agents according to their total income or total wealth, so that the large landowner and the yuppie would be lumped together incorrectly.

We conclude that, for our purpose, we should group agents in the population according to their wage/wealth ratio. We rank all households by their wage/wealth ratio, and find the quintiles of this distribution. The first group contains the families in the lowest quintile of wage/wealth ratio, the second group is the second quintile and so on. Graphically, the split in quintiles is represented by four rays in Figure 1 such that each of the five areas separated by the rays contain 20% of households. The more traditional criterion of classifying families by total income would correspond instead to splitting the sample with four negatively sloped lines, each line representing constant total income. The other traditional criterion of classifying by total wealth would correspond to splitting the sample using four vertical lines.

Another complication stems from the fact that our measures are affected by a pure life cycle effect, something that our model does not take into account. For example, older people are usually wealthier than younger people and they are likely to be retired, which corresponds to $\phi = 0$ in our model. Almost all of them would belong to group 1, thus confusing the life-cycle effect with the wealth effect. We try to remove this effect from our measures by splitting the sample into six age groups, and dividing each age group in five quintiles according to their wage/wealth ratio. The wage of type 1 agents is then calculated with a weighted average of the observed wages of households in the lowest wage/wealth ratio across age groups; the weights given to each age group correspond to percentages of US population as reported by the Census.[16,17]

To summarize, in the benchmark case heterogeneity parameters are obtained by matching each type of agents to the average of each quintile of the distribution of wage/wealth ratios, eliminating the life-cycle effects. In the section on robustness exercises we also calculate the heterogeneity para-

[16] The six age groups are as follows: Less than 25 years old (14.4% of U.S. population), from 25 to 34 (with 23.32% of the population), from 35 to 44 (20.30%), 45 to 54 (13.62%), 55 to 64 (11.43%) and older than 64 (with a 16.89% of total U.S. population).

[17] Recently some work on overlapping generations is available that incorporates a more accurate description of life cycle issues. Some papers in this vein are Conesa, Kitao and Krueger (2007) and Altig and Carlstrom (1999). These papers, however, do not allow for a direct comparison with the Chamley model.

meters splitting the sample with a pure wealth criterion (i.e., splitting the sample by means of vertical lines). The statistics obtained from these two possible criteria are reported in Table 2.

Calibrating the initial wealth of agents in the model with the initial wealth of the quintiles in the data seems problematic, because different assets in the data yield different returns and agents with large wealth may be able to access higher returns. Instead we calibrate initial wealth of agents in the data to the consumption ratio that can be sustained given the actual assets and the actual returns of these assets.

To find the corresponding equilibrium values of λ and the corresponding initial wealth in our model, since the economy is assumed to be in steady state before the reform, individual total income net of taxes is equal to individual consumption each period. For a detailed description on how we compute total capital income see Appendix 1. We calibrate the consumption ratio of agents in the model by finding the ratio of total average labor income plus asset income of different quintiles. The ratios are reported in Table 2. From these consumption ratios we find the initial wealth of each group in the model consistent with steady state and the calibrated consumption ratios in the status quo tax rates. The heterogeneity parameters found in this way and used in the model are reported in Table 3.[18]

3.3 Elasticity of labor

The choice of γ_l is quite important since it governs the elasticity of labor and it will be crucial in determining hours worked after the reform and the impact on welfare of the higher labor taxes.

Ideally we would use a parameter value that matched some basic facts concerning the variability of hours worked. Let us point to two basic well-known facts:

a) *across time* variability of aggregate hours worked is *higher* than variability of aggregate consumption.

[18] As can be seen from Table 3 the consumption ratios that we find can only be sustained if wealth of some of the agents is higher than total capital. This happens because, in the real world, assets such as land play a very important role in the portfolios of individuals, while land is not present in our model. An alternative approach would be to introduce land that delivers returns and services of consumption. Then total wealth would not be larger than productive capital plus land.

b) *across individuals* variability of hours worked is *lower* than variability of consumption.

These observations have been documented by many authors. Fact a) has been emphasized by a number of papers, for example Hansen (1985) and Rogerson (1986). Fact b) has been documented in several contributions and it is confirmed within our calibration of heterogeneity reported in Table 2: the fourth column indicates that agents with the highest number of hours worked (type $j = 2$) work 40% more than type $j = 5$, but they consume almost three times as much. Similar conclusions are derived from the wealth partition.

Fact a) has to do with the reaction of hours worked to a temporal shock to aggregate wealth, while fact b) has to do with the elasticity of hours worked to changes in wealth and wage. The policy experiment that we are considering will cause both a change over time of aggregate hours worked and a redistribution of wealth so that, ideally, we would like to use a model and parameter values that agree with both facts mentioned. Unfortunately, this cannot be done within the standard neoclassical dynamic model.

To see this, we first argue that low values of $|\gamma_l|$ help in explaining fact a), but they are incompatible with fact b). Consider the model with linear utility of leisure, so $\gamma_l = 0$, and assume that agents only differ in their initial wealth, so that $\phi_i = \phi$ for all i. Hansen (1985) and Rogerson (1986) showed that fact a) above can be explained under these assumptions. But (8) implies

$$c_{i,t} = c_{j,t} \qquad \text{for all} \quad t, i, j$$

Therefore, linear utility of leisure implies no volatility of consumption across agents; agents with higher wealth enjoy a higher level of leisure but consume the same as poor agents. This contradicts fact b).

Conversely, we can see that high values of $|\gamma_l|$ fail to explain fact a), but they are compatible with fact b). It is easy to see that in a stochastic model for our choice of B,

$$l_{j,t} \to 1/3 \qquad \text{as} \qquad |\gamma_l| \to \infty,$$

for all j and t. But agents are so averse to changes in hours worked that they are likely to choose low volatility of hours across time and to adapt to fluctuations in income by higher volatility of consumption. Therefore, high values of $|\gamma_l|$ are likely to generate nearly constant hours worked across time in a model with aggregate uncertainty.[19] Hence high $|\gamma_l|$ matches fact b) but it spoils fact a) in a stochastic model.

For our purposes, it seems particularly important to capture how hours worked by agents will react due to a change in capital and labor taxes. For this reason, we choose $\gamma_l = -10$ in the benchmark case which implies a very low wage elasticity of labor. This calibration matches fact b) but it is incompatible with fact a). As with many other parameters, we will check robustness of our main results to this choice.

3.4 Numerical issues

Since before the reform the economy is at the steady state it is trivial to find the equilibrium g. We have to convert the growth model to a model in deviations from trend so as to guarantee that we are solving for variables in a compact set and that the numerical approximations performed can be made, in principle, arbitrarily accurate. Appendix 2 shows how to write the model in terms of deviations from trend.

After the reform, there will be a transition period as allocations converge to the new steady state (in deviations from trend). The numerical difficulty will be, therefore, finding the transition along with the labor tax rate and the ratios λ that will balance the budget constraints after the reform.

Since analytic solutions under the benchmark parameters are not known we resort to numerical simulation. Details on the algorithm we use are given in Appendix 3.

4 Results

To compare with previous results, we first study a homogeneous agent version of our model and we find that suppressing capital taxes causes an improvement in welfare. This confirms that the results in Lucas (1990) and Cooley and Hansen (1992) are replicated for our slightly different model and our calibration. Furthermore, we emphasize the benefits that can be obtained from this reform in a homogeneous agent model relative to the literature, since we find an improvement in welfare even for very high values of risk aversion $-\gamma_c$. We then go on to show the results for the heterogeneous agent case.

[19] To confirm the above claim on variability of hours worked we introduce aggregate uncertainty in section 5.

4.1 Homogeneous agent

4.1.1 Replicating homogeneous agent results

We first consider the effects of suppressing capital taxes in a homogeneous agent version of our economy for the benchmark parameters of Table 1. The values for the variables at steady state are shown in Table 4. The first column shows values for the status quo, while the second column shows the steady state after the reform. As expected the level of capital, labor productivity and even the wage net of taxes are higher in the long run if the reform takes place. The labor tax has to increase from 23% to 37% in order to finance the capital tax cut.

This gain in efficiency in the long run does not translate necessarily in an increase in welfare. As pointed out by Lucas (1990) and Cooley an Hansen (1992), consumption and leisure are lower immediately after the reform (to allow for higher investment and the accumulation of capital), which is a cost of the reform that is ignored in steady state calculations, justifying the concern to calculate the transition explicitly.

The welfare costs of changing the tax system are evaluated, as in previous papers, to be the permanent increase in consumption that would leave each individual indifferent between the status quo and the reform, leaving leisure unchanged. More precisely, letting $\left\{c^A_{j,t}, l^A_{j,t}\right\}$ and $\left\{c^B_{j,t}, l^B_{j,t}\right\}$ be the equilibrium allocations before and after the reform, the welfare gain for agent j is given by π_j that satisfies

$$\sum_t \delta^t \left[u((1+\pi_j/100)\, c^A_{j,t}) + v(l^A_{j,t}, \mu^t)\right] = \sum_t \delta^t \left[u(c^B_{j,t}) + v(l^B_{j,t}, \mu^t)\right].$$

The last line of Table 4 shows that we find a welfare gain for the homogeneous agent of $\pi_H = 5.9\%$. This gain is similar to the one reported in previous papers.

It is not surprising that one obtains an efficiency gain from cutting capital taxes. Status quo capital taxes of 57% are indeed very high. Intuitively this can be seen in figure 2, showing the Laffer curve for the model at hand. The vertical line shows the total revenue from taxes in steady state for each capital tax rate. We notice that the status quo capital tax is close to the top of the Laffer curve, suggesting that it is likely that the economy as an aggregate will benefit from lowering capital taxes.

4.1.2 Emphasizing the efficiency gains of suppressing capital taxes

It has been pointed out that the benefits of suppressing capital taxes in a homogeneous agent model may disappear if the curvature of the utility function with respect to consumption is sufficiently high. To the extent that we are not sure about the true curvature, this brings a word of caution to the efficiency benefits of actually suppressing capital taxes. We reexamine this result and we find that by focussing on models that keep the capital/output ratio constant there is a gain from suppressing capital taxes even for high risk aversion, for a homogeneous agent case. This reinforces the view that suppressing capital taxes is a good policy from the point of view of aggregate efficiency and it will be important for the robustness exercises that we perform in section 5.

The reason that higher curvature in the utility function may limit the benefits of suppressing capital taxes is the following. Increasing $-\gamma_c$ has two effects: first, it causes labor to be more elastic, increasing the costs of a higher labor tax after the reform, second, the initial drop in consumption caused by the cut in capital taxes is more costly if u has more curvature. Indeed, Chari, Christiano and Kehoe (1994) show that if relative risk aversion is $\gamma_c = -8$ suppressing capital taxes would cause a loss in utility in a homogeneous agent case. We find a similar result for our model and calibration in Table 5: even though $\gamma_c = -8$ still shows a small gain in utility, a utility loss is experienced from suppressing capital taxes when $\gamma_c = -11$.

But increasing $-\gamma_c$ and leaving all other parameters constant has some undesirable effects for the calibration of the economy. In order to convert the growth model into a stationary one the effective discount factor becomes $\widetilde{\delta} \equiv \delta\, \mu^{\gamma_c+1}$ (see appendix 2 for details). Therefore, the effective discount factor is lower as $-\gamma_c$ increases and the capital output ratio goes down if all remaining parameters are left unchanged. As shown in Table 5 the steady state capital for $\gamma_c = -11$ is about one fifth of the capital for log utility. This means that for $\gamma_c = -11$ labor at the status quo is much less productive than in the log utility case and it explains why the labor tax rate needs to be raised much more (to 70% instead of 37%) in order to compensate for suppressing capital taxes when risk aversion is high. Therefore changing $-\gamma_c$ relative to the benchmark case not only influences the elasticity of labor and the utility cost of the transition, but it also increases the size of the distortion that labor has to suffer if capital taxes disappear.

We consider more appropriate to analyze the effects of increasing risk aversion in isolation. For this reason we adjust the scaling constant A in the production function so as to maintain the capital output ratio constant while γ_c varies. The results are shown in Table 6. We now find that the homogeneous consumer never looses utility from suppressing capital taxes, even for very high risk aversion.

In summary, the example discussed by Chari, Christiano and Kehoe certainly serves their purpose, namely, to show how ignoring the transition for optimal capital and labor taxes can result in an even lower utility than in the status quo. But it cannot be interpreted as saying that if risk aversion is very high then suppressing capital taxes would be a bad idea, since for reasonable parameter values suppressing capital taxes is always beneficial in terms of aggregate efficiency.

4.2 Heterogeneous agents

The main goal of this paper is to study the welfare effects of eliminating capital taxes when agents are heterogeneous. Since this is a model with complete markets aggregate variables in a heterogeneous agent model behave in a similar way as in the homogeneous agent model of the previous subsection. Therefore, output, investment, capital, gross wages and wages net of taxes increase in steady state under heterogeneity. But under heterogeneous agents abolishing capital taxes also has a redistributive effect. Lower capital taxes mean that a larger part of the tax bill in present discounted terms is paid by those agents with a high wage/wealth ratio. This may offset the gains from the higher aggregate efficiency for these agents. Since we labelled $j = 1$ the agent with the highest relative wealth, a reduction in capital taxes is likely to lower the relative consumption of agent $j = 5$. Therefore, according to equation (12), suppressing capital taxes is likely to increase the ratios $\lambda_j = c_{j,t}/c_{5,t}$ for $j = 1, ..., 4$.

The fact that the reform implies both redistributive and aggregate efficiency effects can be seen from Figure 3, representing the evolution of some variables after the reform. Capital nearly doubles and it is halfway through the new steady state in about 30 quarters. Investment is much higher than in the status quo, as it is even higher in the first few periods than in the new steady state after the reform. Wages increase by about 25%. As expected consumption is very low in the initial periods. Hours worked are higher at the beginning of the transition, showing that the effect of the reform is to induce a higher labor supply. The last two graphs show how consumption and hours worked are very different for agents 1 and 5.

To be more precise about the increased redistribution, Table 7 shows equilibrium ratios of consumption and labor for different capital taxes, in all cases labor taxes adjust to maintain the same level of government spending. The first row corresponds to the status quo capital tax, so it simply describes the equilibrium consumption ratios $\lambda_j = \frac{c_{j,t}}{c_{n,t}}$ and labor ratios before the reform. As expected λ_j is lower for higher j, as we consider agents with a higher wage/wealth ratio. As in the data the cross sectional variation of hours worked is much smaller than the cross-section variation of consumption, justifying our choice of a large $-\gamma_l$.[20]

The last row of Table 7 corresponding to $\tau^k = 0$ shows the effects of suppressing capital taxes. We see that all groups $j = 1, ..., 4$ will consume more and work less, relative to agent 5, after the reform. Furthermore, the one who benefits the most is agent $j = 1$ with the highest relative wealth: while his consumption ratio increases by 70% (it goes from 3.23 before the reform to 5.56) the consumption ratio of agent in the middle quintile $j = 3$ only increases by about 40% (from 2.1 before the reform to 2.94). It is clear, therefore, that lowering capital taxes has a redistributive effect and it lowers the relative consumption of agents with a high wage/wealth ratio such as agents $j = 5$. This shows that the reform redistributes wealth in favor of the agents with a high relative wealth.

The middle rows of table 7 report the effect of four less radical reforms, each reform consisting of cutting the capital tax rate by an additional 20%. That is, the second row shows the equilibrium ratios if capital taxes were lowered permanently to a value of .456, the third row shows the equilibrium if, instead, capital taxes were lowered permanently to a value of .342 and so on. We see the effect is monotone: all λ's increase as capital taxes decrease. These rows will serve to understand the next table.

It is clear from this table that lowering capital taxes increases inequality. But given that there is a gain in aggregate efficiency, as shown in our analysis of the homogeneous agent case, it is possible that even less wealthy agents gain from suppressing capital taxes. To resolve this issue we consider the

[20] Notice, however, that the level of hours worked across agents does not reproduce the data: in the model hours increase with j but they decrease with j in the data. Ideally one would study the effect of suppressing capital taxes with a model that matches these observations, but this would mean going away from the standard neoclassical model so we leave this exercise for future research. The differences of hours worked across agents, in any case, are not large so one would not expect large changes in the main results.

change in welfare for each agent of suppressing capital taxes

Table 8 shows the gains in utility from each of the possible reforms considered in the previous table. If capital taxes were completely suppressed (last row) 40% of the population would be worse off. Perhaps more importantly, agent of type 5 would experience a very large loss in welfare of 32%. Relatively wealthy agents of type 1, on the other hand, benefit greatly from the reform.

We can see that even with a small reduction in capital taxes (first row of table 8) group $j = 5$ with the lowest wage/wealth ratio, would loose welfare, although the rest of the population would benefit.

These welfare comparisons confirm that a policy change that eliminates capital income taxation at the expense of labor income taxation is not Pareto improving. If capital taxes were suppressed, the distributional issues dominate the gain in aggregate efficiency in the sense that they are not Pareto improving and a large part of the population may experience a loss in utility. The loss in welfare for these agents is very high, specially if compared with those reported on the aggregate effects of changes in fiscal or monetary policy using dynamic equilibrium models. We will see in section 5 that these features are very robust to changes in parameter values.

In Table 8 the median voter (agent $j = 3$) does gain from any permanent reduction in capital taxes, but this hardly suggests that suppressing capital taxes at the expense of labor taxes is likely to occur in a modern democracy. First of all because given the very large loss in utility experienced by a large part of the population the reform we consider would be difficult to implement. In modern democracies it is not only the median voter's opinion that matters, as it is difficult to implement a reform in practice if it hurts significantly a sufficiently large part of the population. Second, in the robustness experiments of section 5 we will find that for slightly different parameter values the median voter often looses from suppressing capital taxes. Therefore it is not clear ex-ante that even the median voter will favor such a reform.

4.3 A Pareto improving policy

The results just described indicate that if the long run optimum is implemented from period zero the result is not Pareto improving, as large parts of the population would loose a large amount of welfare. On the other hand the Chamley/Judd result suggests that there must be a way to achieve a pareto-improvement and to lower capital taxes. Are there alternative policy combinations that allow to lower capital taxes and avoid such an increase in inequality?. We have focused so far on the case where only proportional taxes can be levied. A Pareto improvement can be achieved within our model with a lump sum redistribution of wealth in the first period or, equivalently, by choosing an initial value of capital taxes $\tau^k_{j,0}$ different for each agent, so that all agents gain if capital taxes are suppressed.

In this section we study the lump sum transfers in the first period that would achieve a Pareto improvement. Since there are many ways to achieve this Pareto improvement we study the transfers needed to keep constant the ratios λ_i after the tax reform.

It turns out that this policy calls for an expropriation of agents of type 1 to type 4's wealth in the first period; this wealth should then be given to agents of type 5. The size of the expropriation is about 9.4 times the first period's income for $j = 1$ agents, 9.7 the income for $j = 2$, 10 for $j = 3$ and 10.5 for type 4 agents. Type 5 agents obtain then a transfer of about 5.5 times their first period's income and the government uses part of the expropriation to pay for present and future expenditures. The labor tax rate is unchanged and capital taxes are suppressed.

Although this policy does achieve a Pareto improvement, we think this policy is unlikely to be implemented. The redistribution that would be necessary in the first period to improve all agents' welfare would be extremely large: 80% of the population would have to pay nine or ten times their income. This requires a huge amount of credibility on the part of the government, since the above computations assume that agents 1 to 4 are convinced that future capital taxes will be lowered even though taxes are extremely high in the current period. Furthermore it increases enormously the incentive to hide capital income in the first period.

5 Sensitivity Analysis

We discuss the sensitivity of our results to changes in parameter values and empirical calibration. We will introduce uncertainty in the model, this will serve as a robustness check and to explore the empirical validity of the model.

5.1 Parameter Sensitivity

Table 9 shows the welfare gains of all agents from suppressing capital taxes when several parameters of the benchmark case are changed one at a time. In all cases we adjust B so that the hours worked are one third of total time endowment. The column labelled k_{stst} refers to the capital steady state before the reform. The next column shows government spending over output before the reform. Column τ^l contains the labor tax that would operate after the reform. Column π_H indicates welfare improvement in the representative agent version of the model. This can be thought of as a rough measure of the aggregate efficiency gain of suppressing capital taxes for each set of parameters. Columns π_j $j = 1, ..., 5$ show the utility gains of each agent.[21]

We first consider changes in relative risk aversion $-\gamma_c$. Robustness in this dimension is relevant because relative risk aversion is often thought to be larger than one, with values between 2 and 4 much more widely accepted. For each γ_c we adjust the constant A in the production function so as to keep the capital stock constant as explained in section 4.1.2.

Recall that the row for $\gamma_c = -1$ corresponds to the benchmark case. We find that the pattern of gains and losses across agents is similar to the one of the benchmark case but the size of welfare gains or losses is exaggerated by increasing risk aversion.[22] Gains of agents $j = 1, 2$, and losses of agents $j = 4, 5$, are much larger as $-\gamma_c$ increases. Now agent 5 loses 60% of his utility for $\gamma_c = -3$. In addition we find that the median voter $j = 3$ experiences a mild utility loss for reasonable values of relative risk aversion such as 3 or 4. We conclude that for more reasonable values of risk aversion the redistributive effects of suppressing capital taxes are much larger than for log utility and that the median voter will be against the reform for high but not unlikely values of risk aversion.[23]

It is intuitive that higher risk aversion should increase the inequality

[21] In this section we only consider a radical reform of suppressing capital taxes completely, we do not consider the partial reforms of lowering capital taxes.

[22] Only the results up to $\gamma_c = -4$ are reported because the algorithm failed to converge for higher levels of risk aversion. We believe that the algorithm can not find the solution for larger risk aversions because the loss in utility is so large that the algorithm nearly has to divide by zero in various places.

[23] For the case considered in Chari Christiano and Kehoe (1994) where A is constant for all levels of relative risk aversion we obtain even larger welfare losses for low wealth agents. For example, for $\gamma_c = -3$. we find π_1 =64.65%, π_2 =22.12%, π_3 =-3.15%, π_4 =-22.88%, π_5 =-68.49%.

effects of suppressing capital taxes. First of all there is the standard effect of making the initial drop in consumption more costly, which means that the efficiency gain is even lower and there is less to gain from capital taxes. But it is also well known that the wage elasticity of labor is higher for higher risk aversion. This means that for higher $-\gamma_c$ labor goes down more steeply for a given increase in labor taxes and in order to meet the budget constraint the government needs a larger labor tax hike after the reform. As can be seen from Table 9, for a risk aversion of 1 we have $\tau^l = .37$, but for risk aversion of 4 we have $\tau^l = .46$ after the reform. Agents with high wage/wealth ratio have to pay more taxes when risk aversion is higher and they loose relatively more. Also, since labor is more elastic for high risk aversion, the increase in labor taxes is more distortionary and more costly in terms of welfare for the reasons usually explained in public finance taxation.

We also consider robustness to the value of γ_l. As we explained in section 4 the choice for the benchmark case is questionable because it fails to account for the variability of hours worked across time, furthermore, it implies a wage elasticity of about .1 which is lower than usually estimated for the aggregate economy. We see from Table 9 that for lower values of $-\gamma_l$ (and, therefore, closer to those used in the RBC literature) the inequality generated by suppressing capital taxes is only exaggerated. Since lower $-\gamma_l$ implies higher wage elasticity of labor the same discussion as in the previous paragraph justifies the results. Again, for a sufficiently high elasticity the median voter $j = 3$ would now be against the reform.

There is much disagreement about the relevant level of average marginal capital tax rates, so we also study the sensitivity to the tax levels in the status quo. The third panel of Table 9 considers different values for the capital tax before the reform. A lower value for τ^k in status quo causes the redistributive effect to be smaller: agents with high (low) wage/wealth ratio loose (gain) less for lower initial capital taxes. But it is also true that the aggregate gain represented by π_H is smaller if initially the capital tax was not very high. These results are intuitive: if the capital tax is initially low the effects are less strong: the redistributive effect is lower but there is less to be gained from the reform at an aggregate level. The median voter, again, would be marginally against the reform.

Crucial to our results were the heterogeneity parameters determining ϕ and initial wealth of each type of agent. These we calibrated by splitting our sample according to quintiles of the wage/wealth ratio and by removing effects from life cycle. Since this is a relatively non-standard criterion to measure inequality it is worthwhile to explore the effects of the reform using the more traditional criterion of wealth inequality and without adjusting for life cycle. We use the data in the second panel of Table 2 and report the results for this calibration in the fourth panel of Table 9. Again, the large changes in utility are reinforced and the median voter would be against the reform.

5.2 Uncertainty

We end this section by introducing aggregate uncertainty, maintaining the assumption of complete markets. This is useful, first, to confirm robustness of our results on the equity effects of suppressing capital taxes to the introduction of aggregate uncertainty. Second, in section 3.3 we conjectured that our calibrated parameter value for γ_l, chosen to match the low volatility of hours worked across agents, would fail to match the time series volatility of aggregate hours worked. That this is indeed the case can only be studied in a model with aggregate uncertainty as the one we consider here.

We assume that there is a productivity shock multiplying the production function. Therefore, the only change in the technology is that total output at t is now given by $\theta_t F(k_{t-1}, e_t)$ where θ_t is an aggregate productivity shock AR(1) with serial correlation .95 and standard deviation of innovation equal to 0.1. Agents and government can insure against realizations of this shock by accessing a complete market for debt and insurance. The model is carefully laid out in detail in Appendix 4. The only changes in the equilibrium conditions relative to the deterministic case of the main text are that budget constraints (10) hold in expectation given information at $t = 0$ and that the Euler equation (7) holds in expectation given information at t. The model is converted to deviations from trend in a similar way as in the deterministic case.

In order to solve this model we have to find a recursive formulation and iterate on the equilibrium function. An algorithm for this purpose is described in appendix 5.

Table 10 shows some second moments for the model in the status quo that are often considered in real business cycles analysis. As expected, the introduction of heterogeneity maintains the relatively good match for the volatility of investment, output and consumption. As expected from our discussion in section 3.3 time series volatility of hours worked is much lower than in the data. If we assume $\gamma_l = 0$ we would generate much more reasonable volatility of labor across time but the variability of hours worked across agents would be too large.

The welfare results are shown in the last line of Table 9. The welfare losses or gains are quite similar to the case of no uncertainty, the losses and gains slightly larger in absolute value.

As a conclusion to this subsection, our claim that suppressing capital taxes would have very strong redistributive effects is robust to the introduction of aggregate uncertainty.

6 Conclusion

The Chamley/Judd result says that in a model with heterogenous agents and distortionary taxes it is Pareto optimal that capital taxes disappear in the long run. One may interpret this result as saying that all agents are likely to benefit from abolishing capital taxes. We explore whether this is the case in a model with heterogeneous agents. Our model is as close as possible to that of Chamley so as to explore in isolation the effects of heterogeneity.

We find that if capital taxes were suppressed and the lost revenue would be compensated by higher labor taxes the welfare of at least 20% of the population would go down dramatically. For all the experiments we have performed 40% of the population would be worse off. This happens despite the fact that there is always an aggregate efficiency gain from suppressing capital taxes. This result is robust to different parameter values and to the criterion for splitting the sample. For some parameter values, including reasonably high values of relative risk aversion, agents in the lowest quintile of the population loose 60% of their utility.

The effect of suppressing capital taxes on the *median voter* (our type 3 agent) is always quite small. In fact, whether the median voter would gain or loose from the tax reform depends very much on the parameter values chosen for the model. Therefore, from the vantage point of traditional political economy, the model does not give strong predictions about whether such tax reform would be approved in a once-and-for-all referendum. In any case, the loss in welfare for the lowest quintile is so large that it is not surprising that such a reform has not even been considered in actual policy discussions.

Our model is chosen as close as possible to that of Lucas (1990) in order to study the effects of heterogeneity in isolation. We find that there is an aggregate efficiency gain even with high risk aversions, but that in this case the redistributive effect is even larger.

Implementing the long run optimum from period zero would give a dramatically non-pareto improving result and it may not even be favored by the median voter. The abolition of capital taxes would only benefit all agents if there was a very strong reallocation of wealth in the first period. In this sense, the problem of distribution of wealth is several orders of magnitude more important than other traditional topics of macroeconomics. We think that research on distributive and efficiency issues in dynamic equilibrium models is, therefore, a very promising avenue for research.

Capital taxes in the real world are indeed very high, it is probably the case that if capital taxes are lowered this may result in a widespread gain in efficiency. But transferring the burden to labor taxes is unlikely to be implemented in democratic societies, where large minorities have a strong influence in blocking reforms. Dynamic fiscal policy analysis with equilibrium models should help to find ways that capital taxes can be lowered, thereby achieving higher aggregate efficiency, and at the same time insuring that most of the population can benefit from such a reform.

In addressing the calibration of the model we argue that the relevant dimension is not the distribution of total wealth, but the wage/wealth ratio across agents. Agents with a higher wage/wealth ratio experience a large decrease in welfare, while agents with a low wage/wealth ratio would enjoy a large welfare improvement. Therefore the heterogeneity parameters in our model attempt to reproduce the features of the distribution of wage/wealth ratios.

Our intention was to examine the effect of heterogeneity in isolation, therefore we stayed as close as possible to the model of Chamley throughout the paper. Along the way we found a number of empirical issues that this model does not address and that should be resolved in order to examine the effects of reforms in factor taxation. For example, we point that the standard neoclassical model cannot match the observed volatility of hours worked and consumption both across time and across agents. Several modifications of the model may help in resolving this puzzle such as introducing time non-separability in leisure, endogenous human capital accumulation, or the introduction of both an intensive and extensive margin in a model with uninsurable risk. These are left for future research.

Other issues in the calibration of heterogeneity demand a more careful analysis. We treated all families in the same way, but the propensity to consume and work of a family with two children is not the same as that of a single. A better modelling of family issues would be crucial. Also, the model is inconsistent with the observation that agents with higher wage often work more hours, it is likely that a model of human capital and learning by doing will alleviate this problem. Finally, the model has a difficult time explaining total wealth held by all agents and total consumption in a model where all assets yield very similar returns.

All these issues and many others that have been raised in related papers (see references in footnote 4) indicate that there is enormous scope for future research that analyze both the efficiency and distributive effects of changes in the tax code.

APPENDIX 1
Data used in the calibration of the heterogeneity parameters

We have used the Panel Study of Income Dynamics (PSID) to obtain several distributive measures involved in the calibration of the model. This is a well known data set that collects information on families and their offspring. We select families that were interviewed and that kept the same head from 1984 to 1989.

Agents in the model are interpreted as households in the data, not the different individuals that compose each household.

The variables we want to calibrate are the efficiency parameters ϕ_j, and the value of the initial capital stocks $k_{j,-1}$ for each family. For this purpose we look at wages and assets.

The PSID provides measures for average hourly wages, labor income, and several categories of non-human wealth and asset income. These are reported in Figure 1. From these measures we obtain five quintiles in the distribution of $\frac{\phi_j}{k_{j,-1}}$ ratios.

For the actual calibration we need estimate the relative consumption of different groups of agents. For this purpose we comparing the total labor and capital income of different groups and identify the ratio of income to the ratio of consumption.

PSID provides data on labor income. To measure capital income of each family we use the reported measures of asset returns whenever these are available, averaging asset income or rates of return over the last five years of the sample period. Otherwise we multiply each asset's value by average long-run net rate of return as reported in several studies.

In what follows we specify how we find the return of each particular component of non-human wealth.

1. Types of assets for which the PSID reports actual asset returns.

    - Net value of Business or Farms, market and gardening activities , or rooming and boarding activities.
    - Cash assets (savings and checking accounts, CD's, IRA's, etc.) and dividends.

2. Types of assets for which we impute an asset return.

Here we multiply the current value of the asset held by an average (over five years) real rate of return. The following is a list of these assets and the return series we use.

- Net value of Bonds, Insurance Policies and Collectible Goods: Moody's average corporate bond yield[24].
- Stocks, Mutual Funds: S&P'S common stock price index. (Dividends are reported as asset income in the category of 'cash assets').
- Total real estate[25]: we use the value calculated in Rosenthal [1988, p 95]. Rents perceived by the families are already embedded in that rate of return, therefore we do not use the rents reported in the PSID, as to avoid double counting.
- Pensions and Annuities: we use the US Government Security Yield, 10 years or more, Treasury compiled.
- Other Debts: we use the secondary market yields on FHA mortgages since this is composed, mostly, of second mortgages.

We deflate these nominal returns or rates by the wholesale consumer price index. The PSID also reports the net value of autos, mobile homes etc. We do not impute any rent for this category.

APPENDIX 2
The formulation in deviations from trend

We show how the growth model in the paper can be converted into a stationary model by removing growth from the solution. This is a necessary step for obtaining nonlinear approximations to the law of motion. These details are provided for completeness, we follow a standard procedure, only checking that it also works for a model with heterogeneous agents and distorting taxes.

[24] All rates of return or price series were extracted from CITIBANK.

[25] As the difference between real estate value and principal mortgage remaining.

Define deviations from growth be given by

$$\widetilde{c}_{j,t} = c_{j,t}/\mu^t \qquad \widetilde{k}_t = k_t/\mu^t \tag{14}$$

and so on. We want to find equilibrium conditions expressed in terms of deviations from growth; we'll see that the resulting equilibrium conditions can be interpreted as arising from a purely stationary model with the exception of the way depreciation allowances enter the model.

Plugging the expressions (14) in the equilibrium conditions one obtains that feasible allocations satisfy

$$\widetilde{c}_t + g + \widetilde{k}_t - (1-\widetilde{d})\widetilde{k}_{t-1} = \widetilde{F}(\widetilde{k}_{t-1}, \widetilde{e}_t) = A\ \widetilde{k}_{t-1}^{\alpha}\ \widetilde{e}_t^{1-\alpha} \tag{15}$$

satisfies feasibility condition for $\widetilde{d} \equiv 1 - (1-d)/\mu$. Also, $\widetilde{r}_t = \widetilde{F}_k(\widetilde{k}_{t-1},\ \widetilde{e}_t) = r_t/\mu$ and $\widetilde{w}_t = \widetilde{F}_e(\widetilde{k}_{t-1},\ \widetilde{e}_t) = w_t$. Given the utility function (11) hours worked satisfy

$$\widetilde{c}_{j,t}^{\gamma_c}\ \widetilde{w}_t\ (1-\tau^l)\phi_j = B(1-l_{j,t})^{\gamma_l} \tag{16}$$

Therefore, hours worked are stationary and we can take $\widetilde{l}_t = l_t$.

Letting $\widetilde{\delta} \equiv \delta\, \mu^{\gamma_c+1}$ the first order condition for capital can be written as

$$\widetilde{c}_{j,t}^{\gamma_c} = \widetilde{\delta}\ \widetilde{c}_{j,t+1}^{\gamma_c}\ \big((\widetilde{r}_{t+1} - d/\mu)(1-\tau^k) + 1/\mu\big) \tag{17}$$

Notice, here, that $\widetilde{d}$ does not appear in place of the depreciation rate, but d/μ does. It is in this sense that taking into account of growth explicitly modifies the allocations.

The present value budget constraints can be rewritten as

$$\sum_t \widetilde{\delta}^t \left(\frac{\widetilde{c}_{n,t}}{\widetilde{c}_{n,0}}\right)^{\gamma_c} \left[\widetilde{c}_{j,t} - \phi_j \widetilde{w}_t\ l_{j,t}(1-\tau^l)\right] \ = \tag{18}$$

$$k_{j,-1}\mu(1/\mu + (\widetilde{r}_0 - d/\mu)(1-\tau^k)) \qquad \text{for } j = 1, 2, \ldots, n$$

Finally, for the welfare calculations we need the equality

$$\sum_t \widetilde{\delta}^t \left[u(\widetilde{c}_{j,t}) + v(l_{j,t}, 1)\right] = \sum_t \delta^t \left[u(\widetilde{c}_{j,t}) + v(l_{j,t}, \mu^t)\right] \qquad \text{for } j = 1, 2, \ldots, n$$

Notice that in a model without government the growth model would be converted into a stationary model by simply substituting the discount factor δ by $\widetilde{\delta}$, substituting the depreciation rate d by $\widetilde{d}$, and using μ to normalize the returns on capital. But in our case we cannot proceed exactly like this: $\widetilde{d}$ substitutes d in the feasibility constraint but *neither* in the Euler equation of capital *nor* in the budget constraints, where d/μ appears instead. This is because only a portion $1/\mu$ of the depreciation can now be claimed as allowance in the stationary transformation.

APPENDIX 3
Numerical algorithm for the deterministic case

The numerical problem is the following: given τ^k and g, all parameter values and initial conditions, find the deviations from trend $\left\{\widetilde{c}_{n,t}, \widetilde{l}_{n,t}, \widetilde{k}_t\right\}_{t=0}^{\infty}$ and $(\tau^l, \lambda_1, ..., \lambda_{n-1})$ guaranteeing that (16) and (17) hold for $j = n$ and all t, that (15) holds for all t, and that (18) holds for all j, when individual consumption and labor are computed according to (9) for $j = 1, ..., n-1$.

For this purpose we fix large T and consider the following system of equations:

1. (17), (15) and (16) for $t = 0, ..., T-1$.
2. (18).
3. Variables dated $t > T-1$ are set at steady state.

Notice from 1. that we have $3T$ equations and from 2. we have n equations. We have $3T$ unknowns in $\left\{\widetilde{c}_{n,t}, \widetilde{l}_{n,t}, \widetilde{k}_t\right\}_{t=0}^{T-1}$ plus n unknowns in $(\tau^l, \lambda_1, ..., \lambda_{n-1})$. This gives $3T + n$ unknowns and the same number of equations. We know this system of equations cannot be solved exactly, for k_T cannot be at steady state unless the initial capital is at steady state, but the system can be solved approximately by various solution methods and, as $T \to \infty$ we can potentially obtain an arbitrarily accurate approximation. We use $T = 100$ and 150 for robustness.

Notice that since we take the model to be at steady state after T periods the present value budget constraints can be computed exactly.

APPENDIX 4
Characterization of equilibrium

We write down carefully the model under aggregate uncertainty and complete markets. Since the deterministic case is a special case this serves as a careful proof of the discussion in section 2.2 about the sufficient equilibrium conditions.

As usual, we let θ^t denote a possible realization up to period t of the productivity shock and Ω^t the set of all possible realizations θ^t. The consumer chooses $\{x_{jt}\}_{t=0}^{\infty}$ where $x_{jt} \equiv (c_{j,t}, l_{j,t}, k_{j,t})$ where each of these variables are chosen with information available at t, formally, $x_{jt} : \Omega^t \rightarrow R^3$.

Under complete markets it is well known that the budget constraint of the consumer can be summarized in a present value constraint in period zero. Formally, consumer j can be thought of as solving the following maximization problem

$$\max_{\{x_{jt}\}} E_0 \sum_{t=0}^{\infty} \delta^t \left[u(c_{j,t}) + v(l_{j,t}, \mu^t) \right] \tag{19}$$

$$\text{s.t.} \qquad E_0 \sum_{t=0}^{\infty} R_t \left[c_{j,t} - \phi_j \; \mu^t \; w_t \; (1-\tau^l) \; l_{j,t} \right] = k_{j,-1}(1 + (r_0 - d)(1-\tau^k))$$

$$k_{j,-1} \text{ given}$$

taking as given prices $\{R_t, w_t, r_t\}$ where $R_t(\theta^t)$ is the price of one unit of consumption in period t contingent on the realization of shocks up to t is θ^t.

The following proposition reduces the number of equilibrium conditions and simplifies greatly the calculation of equilibrium.

Claim An allocation $\left\{ (c_{j,t}, l_{j,t})_{j=1}^{n}, k_t \right\}$ prices $\{r_t, w_t, R_t\}$ and policy $\left(\tau^k, \tau^l\right)$ are a competitive equilibrium if and only if $\{c_{n,t}, l_{n,t}, k_t\}$ together with some constants $(\lambda_1, ..., \lambda_{n-1})$ satisfy the first order conditions

$$u'(c_{n,t}) = \delta \; E_t \left[u'(c_{n,t+1}) \; \left(1 + (r_{t+1} - d)(1-\tau^k)\right) \right] \tag{20}$$

$$-\frac{v'(l_{n,t}, \mu^t)}{u'(c_{n,t})} = w_t \; (1-\tau^l) \; \mu^t \phi_n \tag{21}$$

and feasibility

$$\frac{1}{n} \sum_{j=1}^{n} c_{j,t} + g_t + k_t - (1-d)k_{t-1} = \theta_t \; F(k_{t-1}, e_t) \tag{22}$$

for all t, and the agents' present value budget constraints

$$E_0 \sum_{t=0}^{\infty} \delta^t \frac{u'(c_{n,t})}{u'(c_{n,0})} \left(c_{j,t} + l_{j,t} \frac{v'(l_{n,t}, \mu^t)}{u'(c_{n,t})} \right) = \tag{23}$$
$$k_{j,-1}(1 + (r_0 - d)(1 - \tau^k)) \quad \text{for} \quad j = 1, 2, \ldots, n$$

hold when individual consumptions and labor are given by (9) for all j.

Notice that the strength of this claim is that the FOC for labor and capital needs to be checked only for one agent. In the above equations e_t is given by (5), factor prices are given by $r_t = F_k(k_{t-1},\ e_t)$ and $w_t = F_e(k_{t-1},\ e_t)$, and

$$R_t = \delta^t \frac{u'(c_{n,t})}{u'(c_{n,0})} \tag{24}$$

Proof

The proof uses standard arguments. The only part that is, perhaps, not obvious is the constancy of the ratio of marginal utilities of consumption (9). This is well known to hold in models without distortions. But (9) follows from the fact that the first order condition of (19) with respect to $c_{j,t}(\theta^t)$ gives

$$\delta^t \frac{u'(c_{j,t}(\theta^t))}{u'(c_{j,0})} = R_t(\theta^t)$$

for all $j = 1, 2, \ldots, n$ all t and all θ^t This implies

$$\frac{u'(c_{j,t}(\theta^t))}{u'(c_{j,0})} = \frac{u'(c_{n,t}(\theta^t))}{u'(c_{n,0})}$$

for all $j = 1, 2, \ldots, n-1$ all t and all θ^t and it implies the first equality in (9), the second equality in (9) follows from (21).

The remaining equations are standard, it is only worth mentioning that we need to keep track of n budget constraints, and the budget constraint of the government can be ignored because it is guaranteed to hold by Walras' law.

That (9), (20), (21) and the budget constraints are sufficient for equilibrium then follows from standard arguments in taxation analysis described, for example, in Chari and Kehoe (1999) and Ljungqvist and Sargent (2004). For example, (9) and (20) imply that the first order condition for capital holds for all agents and so on.

APPENDIX 5
PEA algorithm for the model with aggregate uncertainty

In order to solve numerically for the stochastic model we need to find a policy function that satisfies all equilibrium conditions. Given τ^k and g, all parameter values and initial conditions, the stationary distribution can be solved as follows

- *Step a).* Fix λ_j for $j = 1, 2 \ldots, n-1$ and τ^l.
- *Step b).* Solve for a stochastic process $\left\{\widetilde{c}_{j,t}, \widetilde{l}_{j,t}, \widetilde{k}_t\right\}_{t;j=1,2,\ldots,n}$ that satisfies equilibrium conditions (22), (9), (20) for $j = 1$, and (21) for the values of λ_j and τ^l fixed in the previous step.
- *Step c).* Check if present value constraints (23) are satisfied for the sequence found in Step b). If not, iterate on the above steps until finding values for λ_j and τ^l such that the present value budget constraints are satisfied.

Various algorithms are nowadays available to perform step b) for fixed $\lambda_j, g, \tau^k, \tau^l$. Since there will be a transition to be computed we need a method that can solve for the transition. We now describe how to apply PEA, as described in Marcet and Marshall (1994) or den Haan and Marcet (1990), to perform Step b).

First of all one needs to express the system in terms of deviations from trend, to guarantee that the non-linear approximations are performed on a bounded set and, therefore, have a chance of being accurate. The formulae are analogous to those found in Appendix 2 for the deterministic case.

Given g, τ^k, τ^l and λ's we find $\left\{\widetilde{c}_{n,t}, \widetilde{l}_{n,t}, \widetilde{k}_t\right\}_t$ as follows

- Step B1; Draw a realization θ^T for large T. Here T has to be large enough for the law of large numbers to guarantee that the regression run in the last step is accurate. Substitute the conditional expectation in the right side of (20) by a flexible functional form of the state variables of the model to obtain

$$u'(\widetilde{c}_{n,t}) = \widetilde{\delta}\ \psi(\beta; \widetilde{k}_{t-1}, \theta_t) \tag{25}$$

In practice, ψ is an exponentiated polynomial that is insured to take on only positive values; the parameters β are the parameters in the polynomial. Fix β.

- Step B2. Obtain a long simulation $\left\{\widetilde{c}_{n,t}(\beta), \widetilde{l}_{n,t}(\beta), \widetilde{k}_t(\beta)\right\}_{t=0}^{T}$, consistent with this parameterized expectation. This is done by, in each period, for given state variables, obtaining $\widetilde{c}_{n,t}(\beta)$ from (25), $\widetilde{l}_{n,t}(\beta)$ from (21); finally, $\widetilde{k}_t(\beta)$ is obtained from (22) written in deviations from trend using (9) to find individual consumptions and labor. Given this capital stock we can move to the next period.

- Step B3. Perform a non-linear regression of

$$u'(\widetilde{c}_{n,t+1}(\beta)) \ \left((\widetilde{r}_{t+1}(\beta) - d/\mu)(1-\tau^k) + 1/\mu\right)$$

  (this is the expression inside the conditional expectation in (20) written in deviations from trend) on the functional form

$$\psi(\cdot; \widetilde{k}_{t-1}(\beta), \theta_t).$$

  Call the parameters resulting from this regression $G(\beta)$

- Step B4. Iterate on β to find $\beta_f = G(\beta_f)$.

- The approximate solution is given by $\left\{\widetilde{c}_{n,t}(\beta_f), \widetilde{l}_{n,t}(\beta_f), \widetilde{k}_t(\beta_f)\right\}_{t=0}^{T}$

In order to compute equilibria after the reform, since the initial capital stock is away from the steady state distribution of capital, Step B2 has to be modified. Instead of running one long simulation for large T, run many short run simulations based on independent realizations of the stochastic shock. More precisely, we draw S independent realizations of length T', $\{\theta_{t,s}\}_{t=0,s=1}^{T',S}$ where T' does not have to be as large as before but large enough for $k_{T'}$ to be at the support of the steady state distribution, and S has to be very large. Then, instead of Step B2 we perform

- Step B2'. Obtain simulations $\left\{\widetilde{c}_{n,t,s}(\beta), \widetilde{l}_{n,t,s}(\beta), \widetilde{k}_{t,s}(\beta)\right\}_{t=0,s=1}^{T',S}$, consistent with this parameterized, starting all simulations at fixed initial conditions, and using the steady-state β_f to solve the series at $t = T'$.

This is done, for example, in Marcet and Marimon [1992].

Finally we explain how to evaluate the expectations of infinite sums involved in (23). We draw S realizations of length T'' and approximate the conditional expectation

$$\frac{1}{S}\sum_{s=1}^{S}\sum_{t=0}^{T''}\frac{u'(\widetilde{c}_{n,t,s}(\beta))}{u'(\widetilde{c}_{n,0,s}(\beta))}\delta^{t}\,\left(\widetilde{c}_{j,t,s}(\beta)-\phi_{j}\widetilde{w}_{t,s}(\beta)\,\widetilde{l}_{j,t,s}(\beta)(1-\tau^{l})\right)$$

here, S has to be large so that the law of large numbers guarantees that the above average over realizations is close to the expectation and T'' has to be large so that the infinite discounted sum is well approximated.

The above procedure cuts out the tail of the infinite sum. It certainly works for a high enough T'' but it turns out that in the model at hand an extremely large T'' is needed; the reason is that when capital taxes are lowered, the government accumulates large amounts of debt, and a large primary surplus is needed in the long run. We can obtain accurate solutions for lower T'' if we try to approximate the tail of the above sum, for each s, by the term

$$\frac{\delta^{T''}}{u'(\widetilde{c}_{n,0,s}(\beta))}\,E_{t}\left(\sum_{i=0}^{\infty}\delta^{i}\;u'(\widetilde{c}_{n,t,s}(\beta))\,\left(\widetilde{c}_{j,t+i,s}(\beta)-\phi_{j}\widetilde{w}_{t+i,s}(\beta)\,\widetilde{l}_{j,t+i,s}(\beta)(1-\tau^{l})\right)\right) \tag{26}$$

This approximates the tail of the infinite sum instead of setting it equal to zero. The conditional expectation can be easily approximated by parameterizing the conditional expectation in (26) with a flexible functional form in the steady state, using long run simulations, as is often done in PEA. In this way a much lower T'' can be used. We have found this method to be very efficient and accurate.

REFERENCES

1. Altig D. and C.T. Carstrom (1999) "Marginal Tax Rates and Income Inequality in a Life-Cycle Model", *The American Economic Review*, Vol. 89, No. 5, pp. 1197-1215.

2. Atkeson, A., Chari, V. & Kehoe, P. (1999), 'Taxing capital income: A bad idea', *Federal Reserve Bank of Minneapolis Quarterly Review,* 23(3), 3–17.

3. Carey, David, and Harry Tchilinguirian, "Average Effective Tax Rates on Capital, Labor and Consumption", OECD Economics Department Working Paper no. 258 (2000).

4. Chamley, Christopher (1986). "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives". *Econometrica*, Vol.54, No.3.

5. Chari, V.V., L. J. Christiano and P.J. Kehoe (1994). "Fiscal Policy in a Business Cycle Model", *Journal of Political Economy*, Vol. 102, 617-652,

6. Chari, V.V. and P. Kehoe (1999): "Optimal Fiscal and Monetary Policy" in *Handbook of Macroeconomics,* John Taylor and Mike Woodford, eds. (North Holland: Amsterdam).

7. Conesa J.C., S. Kitao and D. Krueger (2007). "Taxing Capital? Not a Bad Idea After All!", working paper Universitat Autònoma de Barcelona, forthcoming *American Economic Review.*

8. Cooley, Thomas F. and Gary D. Hansen (1992). "Tax Distortions in a Neoclassical Monetary Economy". *Journal of Economic Theory*, Vol.58.

9. Correia I.H. (1999) "On the efficiency and equity trade-off" *Journal of Monetary Economics* 44 (3): 581-603.

10. Domeij D., Heathcote J. (2004). "On the distributional effects of reducing capital taxes" *International Economic Review*, 45 (2): pp 523-554.

11. Feldstein, M.S., L. Dicks-Mireaux and J. Poterba (1983). "The Effective Tax Rate and the Pretax Rate of Return". *Journal of Public Economics*, Vol. 21, 129-158.

12. Flodén, M. (2007). "Why are capital income taxes so high?". working paper IEES Stockholm, forthcoming *Macroeconomic Dynamics.*

13. Greenwood, J., R. Rogerson and R. Wright (1995). "Household Production in Real Business Cycle Theory", *Frontiers of Business Cycle Research*, T. F. Cooley, editor, Princeton University Press.

14. Hansen, Gary D. "Indivisible Labor and the Business Cycle". *Journal of Monetary Economics*, Vol.16 (1985).

15. Joines, Douglass H (1981). "Estimates of Effective Marginal Tax Rates on Factor Incomes". *Journal of Business*, vol.54, no.2, 191-226.

16. Jones, L., R. Manuelli and P. Rossi (1993). "Optimal Taxation in Models of Endogenous Growth". *Journal of Political Economy*, Vol.101, No.3.

17. Judd, Kenneth L. (1985). "Redistributive Taxation in a Simple Perfect Foresight Model". *Journal of Public Economics*, Vol. 28, 59-83.

18. Judd, Kenneth L. (1987). "The Welfare Cost of Factor Taxation in a Perfect Foresight Model". *Journal of Political Economy*, Vol.95.

19. Klein P, P. Krusell and V. Rios-Rull (2007) "Time Consistent Fiscal Poicy", working paper, U. Penn. Forthcoming *Review of Economic Studies.*

20. L. Ljungqvist and T.J. Sargent (2004), *Recursive Macroeconomic Theory*, 2nd edition, MIT press.

21. Lucas, Robert E. Jr. (1990). "Supply-Side of Economics: An Empirical Review". *Oxford Economic Papers,* Vol. 42, pp 293-316.

22. Maliar L. and S. Maliar (2001). "Heterogeneity in Capital and Skills in a Neoclassical Stochastic Growth Model", *Journal of Economic Dynamics and Control*, 25/9, pp. 1367-1397.

23. Marcet, Albert, and Ramon Marimon (1992). "Communication, Commitment and Growth". *Journal of Economic Theory*, Vol.58, No.2.

24. Marcet, Albert, and David A. Marshall (1994). "Solving Non-linear Rational Expectations Models by Parameterized Expectations; Convergence to Stationary Solutions". Working Paper, Department of Finance, Northwestern University and Universitat Pompeu Fabra.

25. McGrattan, E., R. Rogerson and R. Wright. (1997) "An Equilibrium Model of the Business Cycle with Household Production and Fiscal Policy".*International Economic Review*, Vol. 38 (2). 267-90.

26. Mendoza, E., A. Razin and L. Tesar (1994) "Effective Tax Rates in Macroeconomics. Cross Country Estimates of Tax Rates on Factor Incomes and Consumption", *Journal of Monetary Economics, Vol.* 34, Issue 3, 297-323.

27. Rogerson, R. (1986). "Indivisible Labor, Lotteries and Equilibrium". *Journal of Monetary Economics*, Vol.21.

28. Rosenthal, Stuart S. (1988). "A Residence Time Model of Housing Markets". *Journal of Public Economics*, Vol. 36, 87-109.

Table 1: **Benchmark Calibration. Technology, utility and policy parameters.**

α	.36	τ^l	.23
δ	.99	τ^k	.57
d	.02	μ	1.004
γ_c	-1.	$k_{g,-1}$	-2.0
γ_l	-10.	A	1

Table 2: **Means and ratios by quintiles, PSID sample**

Wage/Wealth partition						
	Means by type			Ratios of type i over type 5		
Type	Hours	Wage	Income	Hours	Wage	Income
1	2708.03	7.89	58611.94	1.315	1.048	3.241
2	2837.86	11.11	50397.86	1.378	1.475	2.787
3	2468.28	9.72	37822.32	1.199	1.291	2.092
4	2333.49	9.4	31790.4	1.133	1.248	1.758
5	2059.41	7.53	18083.11	-	-	-
Wealth partition						
1	3031.43	15.04	84644.67	1.597	2.549	5.708
2	2858.14	10.31	45058.34	1.505	1.747	3.039
3	2520.16	7.99	31277.28	1.327	1.354	2.109
4	2098.94	6.48	21047.11	1.106	1.098	1.419
5	1898.61	5.9	14828.54	-	-	-

Type 1 corresponds to households with a lower wage/wealth ratio or a lower wealth

Table 3: **Heterogeneity parameters. Benchmark Economy.**

Wage/Wealth Partition		*Wealth Partition*	
ϕ_1/ϕ_5	1.05	ϕ_1/ϕ_5	2.55
ϕ_2/ϕ_5	1.48	ϕ_2/ϕ_5	1.75
ϕ_3/ϕ_5	1.29	ϕ_3/ϕ_5	1.35
ϕ_4/ϕ_5	1.25	ϕ_4/ϕ_5	1.10
$k_{1,-1}/k_{-1}$	5.54	$k_{1,-1}/k_{-1}$	10.39
$k_{2,-1}/k_{-1}$	1.76	$k_{2,-1}/k_{-1}$	0.87
$k_{3,-1}/k_{-1}$	0.35	$k_{3,-1}/k_{-1}$	-0.85
$k_{4,-1}/k_{-1}$	-0.63	$k_{4,-1}/k_{-1}$	-2.76

Table 4: **Steady state, homogeneous agent, before and after reform**

Variable	Status Quo	Zero capital Tax
τ^k	0.57	0
k	6.72	13.21
$invest$	0.16	0.32
GNP	0.98	1.25
l	0.333	0.331
c	0.57	0.68
w	1.89	2.41
r	0.05	0.03
τ^l	0.23	0.37
$w(1-\tau^l)$	1.46	1.52
g	0.25	0.25
π_H		5,90%

Table 5: **Utility gain from suppressing capital taxes, homogeneous agent, increasing $-\gamma_c$.**

$-\gamma_c$	k_{stst}	g	τ^l	π_H
0.5	7.77	0.26	0.35	6.34%
1	6.72	0.25	0.37	5.90%
3	4.17	0.21	0.45	4.25%
5	2.88	0.17	0.51	2.97%
8	1.87	0.12	0.61	1.39%
11	1.33	0.08	0.70	-0.17%

The first column refers to the parameter varied. Columns 2 - 5 indicate how the calibration and results change for the homogeneous agent case. τ^l is the labor tax rate after suppressing capital taxes in this case, while π_H measures the welfare gain when agents are homogeneous.

Table 6: **Utility gain from suppressing capital taxes, homogeneous agent, increasing γ_c, keeping K/L constant**

$-\gamma_c$	k_{stst}	g/y	τ^l	π_H
0.5	6.72	0.25	0.35	6.31%
1	6.72	0.25	0.37	5.90%
3	6.72	0.27	0.44	4.52%
4	6.72	0.27	0.46	4.05%
5	6.72	0.28	0.47	3.69%
8	6.72	0.28	0.50	3.06%
11	6.72	0.29	0.52	2.71%
14	6.72	0.29	0.53	2.20%
18	6.72	0.30	0.54	0%
22	6.72	0.30	0.55	0%

The first column refers to the parameter varied. Columns 2 - 5 indicate how the calibration and results change for the homogeneous agent case. τ^l is the labor tax rate after suppressing capital taxes in this case, while π_H measures the welfare gain when agents are homogeneous.

Table 7: **Consumption and labor ratios**

New τ^k	*Wage/Wealth Partition*							
	$\frac{c_1}{c_5}$	$\frac{c_2}{c_5}$	$\frac{c_3}{c_5}$	$\frac{c_4}{c_5}$	$\frac{l_1}{l_5}$	$\frac{l_2}{l_5}$	$\frac{l_3}{l_5}$	$\frac{l_4}{l_5}$
0.57	3.23	2.77	2.10	1.77	0.95	0.94	0.92	0.89
0.456	3.57	3.00	2.21	1.82	0.95	0.93	0.92	0.88
0.342	3.85	3.11	2.31	1.88	0.95	0.93	0.91	0.88
0.228	4.34	3.43	2.47	2.00	0.94	0.93	0.91	0.87
0.114	4.76	3.67	2.62	2.10	0.94	0.92	0.90	0.86
0	5.56	4.11	2.94	2.28	0.94	0.92	0.90	0.85

Table 8: **Welfare gains in benchmark case**

New τ^k	π_1	π_2	π_3	π_4	π_5
0.456	6.67%	3.56%	2.22%	0.70%	-4.05%
0.342	12.38%	5.88%	3.08%	-0.07%	-9.90%
0.228	17.52%	7.44%	3.08%	-1.79%	-16.86%
0.114	22.33%	8.50%	2.54%	-4.13%	-24.51%
0	26.98%	9.26%	1.62%	-6.89%	-32.60%

Table 9: **Sensitivity analysis: Effects of parameter variations on calibration and welfare gains of fully suppressing capital taxes**

	Risk aversion								
$-\gamma_c$	k_{stst}	g/y	τ^l	π_H	π_1	π_2	π_3	π_4	π_5
0.5	6.72	0.25	0.35	6.31%	20.27%	7.94%	3.01%	-1.92%	-18.56%
1	6.72	0.25	0.37	5.90%	26.98%	9.26%	1.62%	-6.89%	-32.60%
3	6.72	0.27	0.44	4.52%	51.09%	17.19%	-2.53%	-19.18%	-60.48%
4	6.72	0.27	0.46	4.05%	73.28%	22.12%	-3.77%	-22.88%	-66.64%
	Labor disutility								
$-\gamma_l$	k_{stst}	g	τ^l	π_H	π_1	π_2	π_3	π_4	π_5
15	6.72	0.25	0.37	6.05%	26.12%	9.03%	1.74%	-6.36%	-30.98%
10	6.72	0.25	0.37	5.90%	26.98%	9.26%	1.62%	-6.89%	-32.60%
1	6.72	0.25	0.38	4.32%	57.07%	9.72%	-6.82%	-23.15%	-61.58%
	Status Quo capital taxation								
τ^k	k_{stst}	g	τ^l	π_H	π_1	π_2	π_3	π_4	π_5
40	9.09	0.23	0.33	1.74%	12.58%	3.41%	-0.55%	-4.99%	-18.74%
30	10.31	0.21	0.30	0.74%	7.78%	1.81%	-0.78%	-3.68%	-12.75%
20	11.41	0.20	0.27	0.24%	4.39%	0.86%	-0.67%	-2.39%	-7.79%
	Wealth Partition								
	6.72	0.25	0.37	5.9%	36.91%	5.48%	-8.08%	-37.67%	-49.38%
	Uncertainty								
	6.77	0.25	0.37	5.9%	28.36%	10.35%	1.31%	-7.84%	-36.10%

The first column refers to the parameter varied. Columns 2 - 5 indicate how the calibration and results change for the homogeneous agent case. τ^l is the labor tax rate after suppressing capital taxes in this case, while π_H measures the welfare gain when agents are homogeneous.

Table 10: **First and second moments of aggregate variables Stochastic Status Quo**

Variable	Mean	Std. Error	Volatility
k	6.77	0.47	0.07
i	0.16	0.02	0.13
GNP	0.99	0.05	0.05
l	0.33	0.0009	0.0027
c	0.58	0.03	0.05

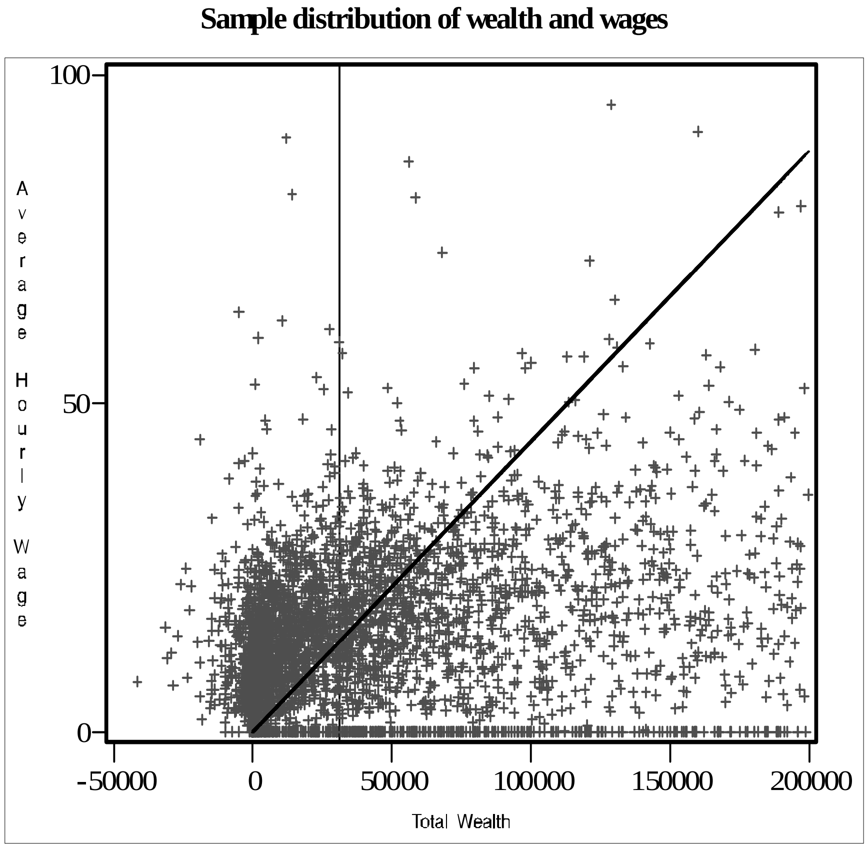


Figure 1: The wealth and wage ranges have been chosen for a better graphical representation of the diversity of wage/wealth ratios. These ranges leave out 12% of the sample. The positively sloped line shows how the sample would be split in two parts according to the wage/wealth criterion. The vertical line shows the split in two parts according to wealth.

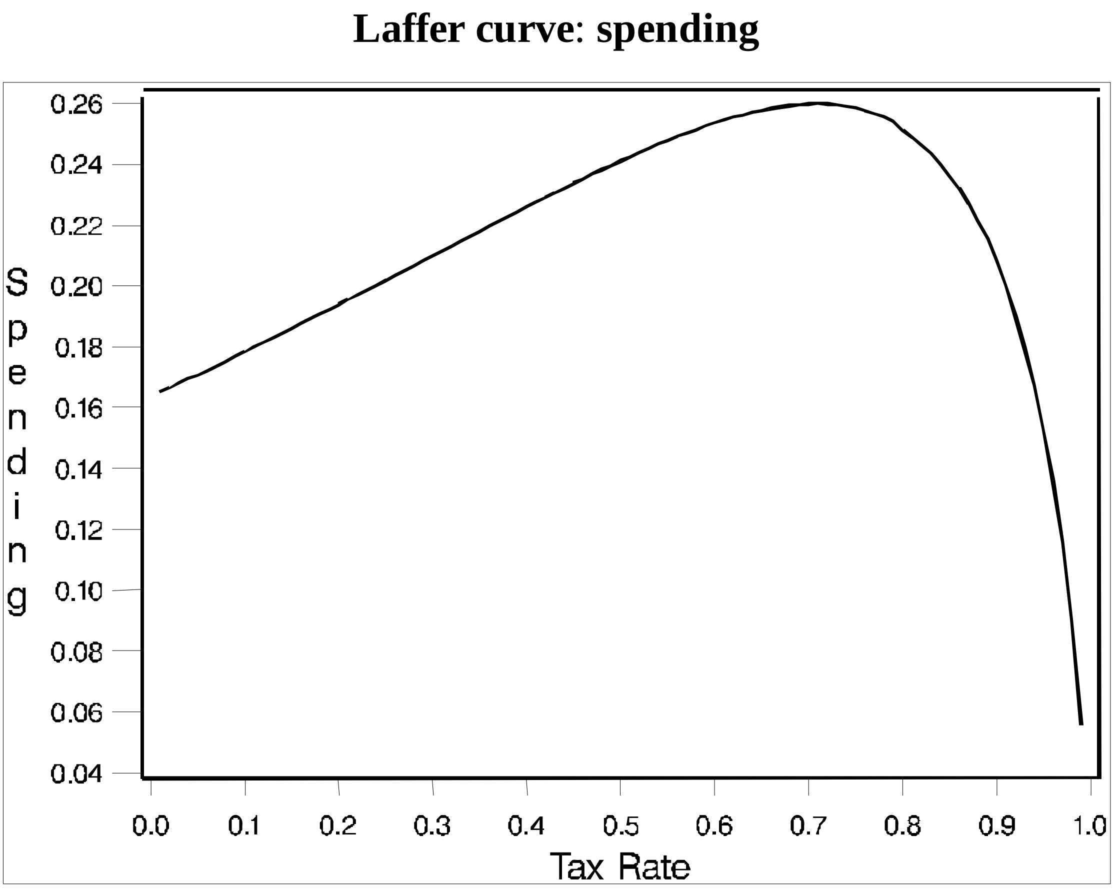
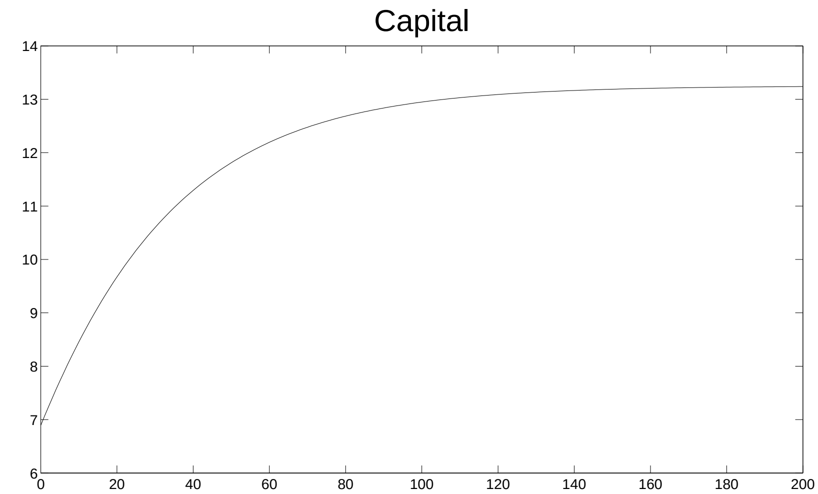
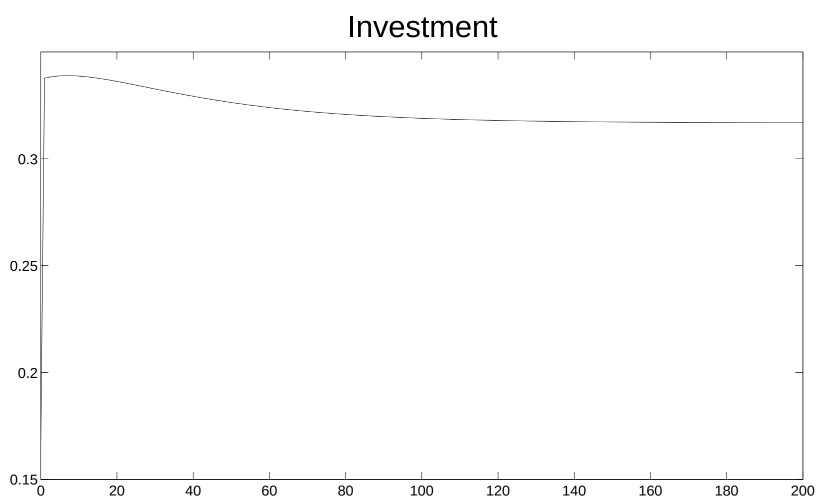

Laffer curve: spending
Spending
0.26
0.24
0.22
0.20
0.18
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
Tax Rate

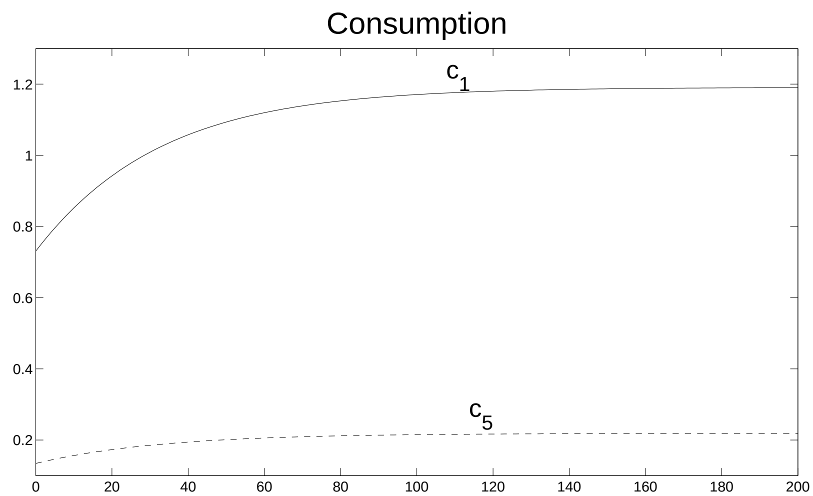
Figure 2:

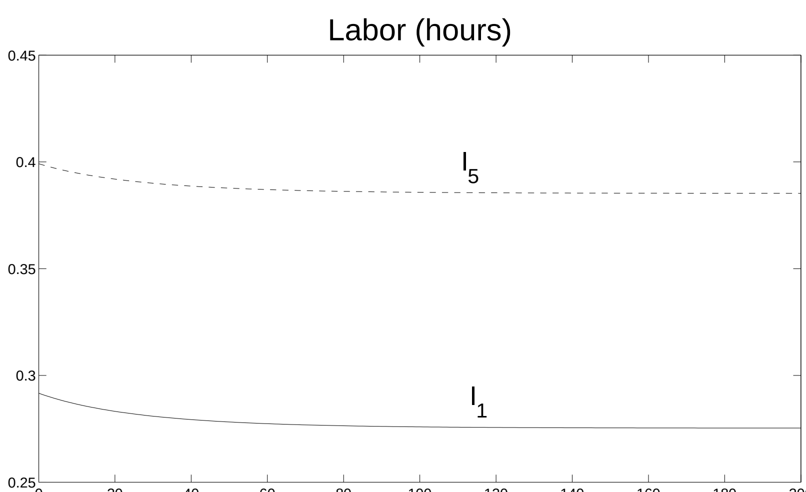
Figure 3


Capital
14
13
12
11
10
9
8
7
6
0
20
40
60
80
100
120
140
160
180
200


Investment
0.3
0.25
0.2
0.15
0
20
40
60
80
100
120
140
160
180
200


Wages
2.5
2.4
2.3
2.2
2.1
2
1.9
1.8
0
20
40
60
80
100
120
140
160
180
200


Consumption
c1
c5
1.2
1
0.8
0.6
0.4
0.2
0
20
40
60
80
100
120
140
160
180
200


Labor (hours)
l5
l1
0.45
0.4
0.35
0.3
0.25
0
20
40
60
80
100
120
140
160
180
200