THE
ECONOMIC
JOURNAL

The Economic Journal, **120** (*March*), 105–130. doi: 10.1111/j.1468-0297.2009.02333.x. © The Author(s). Journal compilation © Royal Economic Society 2009. Published by Blackwell Publishing, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA.

SUPPLY SIDE INTERVENTIONS AND REDISTRIBUTION*

Teresa Garcia-Milà, Albert Marcet and Eva Ventura

We evaluate the effect on welfare of shifting the burden of capital income taxes to labour 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: it increases efficiency in terms of aggregate production and it redistributes wealth in favour 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.

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 just 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 economic success of Ireland is often linked to lower capital taxes. Most empirical measures of capital taxes show that these 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 labour, with an average annual decrease of −0.2% in capital taxes, and an increase of 0.3% on labour taxes. For the US these rates are −0.5% and 0.2% 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 *et al.* (1993)

* We thank the editor, A. Scott, and two anonymous referees. N. Guner, V. Ríos-Rull, J. Geweke, E. McGrattan, A. Braun, E. Prescott, L. Christiano, O. Attanasio, M. Reiter, N. Roubini, D. Epple and C. Sims provided helpful suggestions. 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 Ministry of Science and Technology of Spain grants SEJ2007-64340, SEJ2006-13537, Instituto de Estudios Fiscales, CREI and Barcelona GSE Research Network and of the Government of Catalonia. All errors are our own.

[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, 1987). A proof for the model considered in the current article where no lump sum transfers are available is found in Atkeson *et al.* (1999).

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 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 labour 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 labour 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 discussing capital vs. labour taxes. The object of the current article 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 labour 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 ratio 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 lose from the reform.

Some papers have shown how it may be difficult to implement Ramsey policies due to time inconsistency. For example, Klein *et al.* (2008) show how a time consistent policy under balanced budget 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 say the new constitution would have to be written very carefully in order to be approved, since it would have to implement the actual transition of optimal taxes under full commitment. The median voter is likely to disagree with a change in the constitution stating that capital taxes are immediately suppressed, and a significant part of the population would very strongly disagree.

Since we are extrapolating the behaviour of the economy into an area where no observations are available, the answer we find is highly dependent on both theoretical and empirical elements introduced in the analysis. In the article we provide a careful discussion of how to capture the features of the joint distribution of labour and capital income across agents that are relevant for the exercise at hand. Also, we discuss carefully the effects of different assumptions on the elasticity of labour. 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 the joint distribution of wealth and labour income across the population is what matters and, in particular, that the relevant dimension of this distribution for us is the dispersion of the wage/wealth ratio across agents.[3] Our approach is to match the observed distribution of the wage/wealth ratio. Another key aspect in the calibration is the parameter values and functional forms that relate to the elasticity of labour, since these will influence the efficiency cost of the higher labour taxes that are needed to compensate for the lost capital tax revenue. We argue that the standard neoclassical model does not allow us 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 labour supply to roughly match the cross-section observations.

This is a revised version of our working paper that was first circulated in 1995.[4] Other papers have analysed related issues since our first working paper came out. Correia (1999) shows the source of redistributive effects analytically in a model with aggregation, Domeij and Heathcote (2004) use a model with incomplete markets and focus on the effects of idiosyncratic uncertainty.[5] Flodén (2009) studies a model where the transition of capital taxes is optimal from the point of view of 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 with those of our 1995 working paper. Conesa *et al.* (2009) find that in a model with overlapping generations and idiosyncratic uninsurable risk it is often optimal *ex ante* to have high capital taxes. A summary of these papers is that our main finding is very robust: suppressing capital taxes has large redistributive effects that would strongly decrease the welfare of large parts of the population under many extensions of our model. Our article still is the closest one to Chamley's framework, so it shows the effect of heterogeneity in isolation and in a simple model.

Some available work has used models where aggregation obtains. For example, Correia (1999), Domeij and Heathcote (2004) and Flodén (2009) use Greenwood Herkowitz and Huffman (GHH) preferences. In those cases one can solve first for the aggregate solution and then disaggregate the results. But in a model with growth these preferences imply zero hours worked in steady state and, therefore, in the *status quo* economy. GHH also presents some problems in matching volatility of hours. Therefore we use a model without aggregation and solve explicitly for the disaggregated choices

[3] Few papers have stressed the importance of the *joint* distribution of wealth and wage earnings. Krusell and Ríos-Rull (1999) note how results in a model of political economy are sensitive to whether consumers are ranked according to wealth distribution or to earnings distribution. Domeij and Heathcote (2004) also discuss the correlation of earnings and wealth across agents in the data.

[4] Some differences with that version are that we have now five agents instead of two, we now only consider a deterministic model, there are many more robustness checks and we have added the analysis for the high risk aversion case.

[5] The 'no-earnings-risk' economy of Domeij and Heathcote (2004) amounts to redoing our exercise for GHH preferences and without growth.

of each type of agent. However, given our approach this only increases the computation costs mildly relative to a homogeneous agent model. Since under GHH preferences one has to resort to numerical solutions for the aggregate variables anyway, the increase in computational cost from having a model without aggregation is quite minor.

Along the way, we re-examine the result of Chari *et al.* (1994) that suppressing capital taxes would be undesirable in a model with a representative agent and high risk aversion. We find that if the calibration maintains a roughly plausible capital output ratio, suppressing capital taxes is beneficial for a representative agent even with high risk aversion. However, in a model with heterogeneous agents and high risk aversion, the redistributive effects of suppressing capital taxes are even higher.

The layout of the article is as follows. The model is presented in Section 1. Section 2 discusses issues pertaining to parameter calibration using data from the US economy. Section 3 presents the results derived from the simulations. Section 4 performs sensitivity analysis. The conclusion ends the main article. Appendix A discusses the details of the calibration using PSID data set and Appendix B discusses the numerical algorithm in detail.

1. The Model

In this Section we describe a simple neoclassical growth model with heterogeneous agents, endogenous production, labour choice, exogenous deterministic growth[6] and government spending. Government can only use distortionary capital and labour taxes. Agents differ both in terms of their human, and non-human wealth.

1.1. *Consumer, Firm, and Government Behaviour*

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 normalised 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 the productivity of their endowment of labour hours and initial capital stock. Income from labour and capital are taxed at constant rates τ^l and τ^k.

Consumers of type j solve the following maximisation problem:

$$
\begin{aligned}
&\max_{\{x_{jt}\}} \sum_{t=0}^{\infty} \delta^t \left[u(c_{j,t}) + v(l_{j,t}, \mu^t)\right] \\
&\text{s.t. } c_{j,t} + k_{j,t} = \phi_j \mu^t w_t l_{j,t}(1-\tau^l) + k_{j,t-1}\left[1 + (r_t - d)(1-\tau^k)\right] \\
&\text{given } k_{j,-1}
\end{aligned}
\tag{1}
$$

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

[6] Introducing growth explicitly is important in order to quantify the effect 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 analysis was based on the stationary version of the model. This is made explicit in Appendix B where we show the equations for the model in deviations from trend.

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 denotes the wage for efficiency units of work, r_t capital rental, these prices normalised in terms of the consumption good of the period. The wage obtained per hour worked in period t by agent j is $\phi_j \mu^t w_t$, reflecting the fact that this agent produces ϕ_j efficiency units per hour worked and that labour productivity grows exogenously at the rate μ.[7] Since we concentrate our study on issues of distribution, our agents only differ in their initial wealth $k_{j,-1}$ and their efficiency of labour $\phi_j > 0$; these are normalised 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 only the capital income net of depreciation allowances is taxable.

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. Individual capital holdings could be negative if the agent holds some debt.[8]

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

$$\begin{aligned} &\max_{(y_t, e_t, k_{t-1})} y_t - w_t e_t - r_t k_{t-1} \\ &\text{s.t.} \quad y_t = F(k_{t-1}, e_t) \end{aligned} \tag{2}$$

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

Since total supply of efficiency units of labour is $\frac{1}{n}\mu^t \sum_{j=1}^{n} \phi_j l_{j,t}$ all variables grow at a rate μ in the steady state except labour, which is constant in steady state. Normalising the group size to $1/n$ together with $\frac{1}{n}\sum_{j=1}^{n} \phi_j = 1$ 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 and grows at the same rate as output, so the sequence of government consumption is given by $g_t \equiv \mu^t g$ for a given constant g.[9] Tax revenues accrue from constant capital and labour tax rates τ^k, τ^l. Government can save or dissave by borrowing or lending at equilibrium interest rates. As is well known this is equivalent

[7] Introducing the trend of labour 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 of consumption, where the term μ^t drops out. We only need the term μ^t in the utility function for the high risk aversion cases considered in the robustness exercises in Section 4.

[8] 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.

[9] Since we maintain g constant across policy experiments, the equilibrium computed and the welfare gains discussed in Sections 3 and 4 are consistent with a model where government spending enters the utility function or the production function. To keep notation simple, we write the article as if government spending has no productive use.

with assuming that the government has (possibly negative) capital stock holdings k_t^g. This amounts to the following budget constraint at period-t

$$g_t + k_t^g = \tau^k (r_t - d) k_{t-1} + \tau^l w_t e_t + [1 + (r_t - d)(1 - \tau^k)] k_{t-1}^g \tag{3}$$

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

1.2. *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 maximise utility and firms maximise 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, labour and consumption good are given, for all t, by

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

$$\frac{1}{n} \mu^t \sum_{j=1}^{n} \phi_j l_{j,t} = e_t \tag{5}$$

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

For interior solutions, the first order conditions for capital and labour in the consumer's problem are

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

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

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

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

It is easy to see that these equilibrium conditions can be summarised in the following way. Equation (7) implies that for some constants λ_j

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

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.

Substituting (7) and (8), and substituting for individual savings in the consumer budget constraint we obtain the present value formulation of the consumers' budget constraints

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

The budget constraint of the government is guaranteed by Walras' law and, therefore, can be ignored.

It is easy to see that given a policy (g, τ^k, τ^l) necessary and sufficient conditions for $\{(c_{j,t}, l_{j,t})_{j=1}^n, k_t\}_{t=0}^{\infty}$ to be an equilibrium sequence are[10]

1 for all $t = 0,1,\ldots$ the following equations hold: (6), (7) for $j = n$, (8) for $j = n$, and (9) for some $\lambda_1,\ldots,\lambda_{n-1}$.
2 (10) for $j = 1,\ldots,n$

This reduces the number of variables and equations that need to be found to compute an equilibrium, since (7) and (8) for $j = 2,\ldots,n$, period-t budget constraints (1) and (3) can be ignored. Notice that the way we formulate the problem involves finding the individual variables directly, we do not use any aggregation result, as there is no aggregation result that holds for this model. An algorithm is described in detail in Appendix B which in this model implies negligible increase in computational costs due to heterogeneity.

2. Calibration, Stylised Facts, Analytic Results and an Algorithm

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

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

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 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 behaviour of actual economies in the dimensions that are relevant for our exercise. We now describe the criteria that guided our choice of parameter values in the benchmark economy.

[10] For details see Appendix 5 of the 1995 working paper version. That paper presents the case with uncertainty which encompasses the certainty case.

2.1. *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. The values we use are summarised in Table 1.

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:

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

As usual, B is chosen so that the representative agent works 1/3 of his time endowment in the steady state corresponding to the *status quo*. Also, α is chosen to match the labour share of income. Depreciation rate, discount rate of utility, parameter A and growth rate are set to the usual values for quarterly data.

As for policy parameters (τ^l, τ^k, g), tax rates 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 *et al.* (1997) estimates of $\tau^k = 0.57$ and $\tau^l = 0.23$ 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 *et al.* (1994) methodology, obtain estimates of around 0.5 for capital tax and 0.22 for labour tax for the period 1980-97.[11] We discuss in detail the sensitivity of our results to the value of τ^k.

Government spending g is selected to balance the government budget constraint in *status quo* steady state.[12]

Initial aggregate capital is set at the steady state of the *status quo* policy.[13]

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 per cent in the *status quo*.

[11] The rate of $\tau_k = 0.57$ 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 labour and income taxes, specially on the level of the capital tax. Feldstein *et al.* (1983) obtain estimates of τ^k that range between 0.55 and 0.85 for the period 1953–79. Cooley and Hansen use a lower tax rate, setting $\tau^k = 0.5$ (this number is based on Joines (1981) with the data ending in 1979) and they do not subtract growth from the depreciation allowances; Chari *et al.* use $\tau^k = 0.27$; Lucas (1990) considers capital and labour taxes of 0.4; Greenwood *et al.* (1995) set $\tau^k = 0.70$.

[12] Since we are interested in the effects of substituting capital taxes by labour 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.

[13] Table 4 shows the values of capital and output. The capital/output ratio in status quo is about seven, lower than the values of ten or twelve that are often used 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.

Table 1

Benchmark Calibration. Technology, Utility and Policy Parameters

α	0.36	τ^l	0.23
δ	0.99	τ^k	0.57
d	0.02	μ	1.004
γ_c	−1.0	$k_{g,-1}$	−2.0
γ_l	−10.0	A	1

2.2. *Heterogeneity Parameters*

The parameters that determine agents' heterogeneity, namely the productivity of labour ϕ_j and initial levels of wealth $k_{j,-1}$, are key to the outcome of the policy reform under study. 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. Two agents with the same wage/wealth ratio are likely to be affected in the same way by a tax reform, even if one of them has a much higher total income than the other. The following concrete example demonstrates this point. Consider the case where the wage/wealth ratio is constant across all agents:

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

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

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

In other 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 ϕ_j/ϕ_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 twice that of agent j before the reform, agent i will continue to consume twice that of agent 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 lose from suppressing capital taxes. Therefore, we should examine if (13) is a good approximation to the empirical distribution of income.

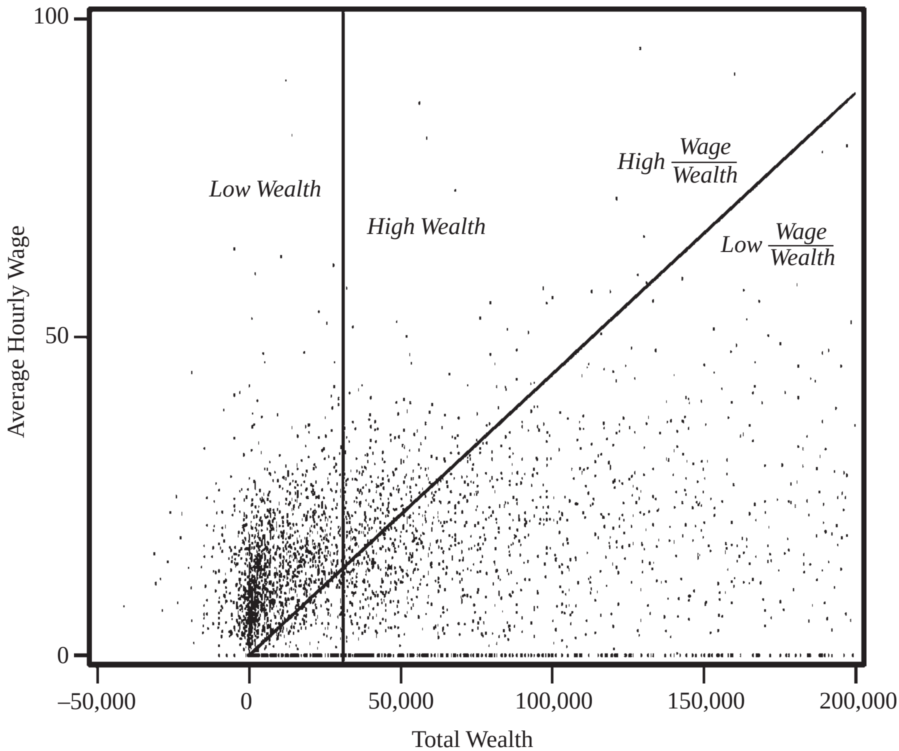

Fig. 1. *Sample Wages and Wealth*

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.[14] 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 lose from the abolition of capital taxes because most of their income comes from labour, which will be taxed more heavily after the reform. Agents with a similar wage/wealth ratio either all gain or all lose, regardless of their total wealth.

To give some names to the situation: it is not that important for us to distinguish between a very highly qualified and a low qualified worker if their levels of wealth are both low. These workers might have a very different level of income but both have a high wage/wealth ratio. It is important, however, to distinguish between a very highly

[14] The details on how this Figure has been constructed are in Appendix A.

qualified worker and a large landowner who has zero labour income: they both have a high total income but they have very different wage/wealth ratios. 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 highly qualified worker would be lumped together incorrectly, because the first is likely to gain from the reform we consider while the latter is likely to lose.

In order to capture the observed distribution of wealth/wage ratios we rank all households by their wage/wealth ratio and find the quintiles of this distribution. Each type in the model will represent one of the quintiles. Graphically, the split in quintiles would be represented by four rays in Figure 1 such that each of the five areas separated by the rays contains 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 a given level of 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 into 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.[15,16]

To summarise, in the benchmark case heterogeneity parameters ϕ_j, $k_{j,-1}$ are obtained by matching each type of agents in the model 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 parameters 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 are often able to access higher returns. Instead we calibrate λ to the ratio of consumption that can be sustained by total labour and capital income of each agent given the actual assets and the actual returns of these assets for the agents in the sample. For a detailed description on how we compute total capital income see Appendix A. 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

[15] The six age groups are as follows: less than 25 years old (14.40% of the US population), from 25 to 34 (23.32%), from 35 to 44 (20.30%), 45 to 54 (13.62%), 55 to 64 (11.43%) and older than 64 (16.89%).

[16] Conesa *et al.*, (2009) explore the effect of capital taxes in an overlapping generations model. Therefore, they are better able to match income through the life cycle.

Table 2

Means and Ratios by Quintiles, PSID Sample

Type	Hours	Wage	Income	Hours	Wage	Income
	Means by type			Ratios of type i over type 5		
Wage/Wealth partition						
1	2,708.03	7.89	58,611.94	1.315	1.048	3.241
2	2,837.86	11.11	50,397.86	1.378	1.475	2.787
3	2,468.28	9.72	37,822.32	1.199	1.291	2.092
4	2,333.49	9.4	31,790.4	1.133	1.248	1.758
5	2,059.41	7.53	18,083.11	–	–	–
Wealth partition						
1	3,031.43	15.04	84,644.67	1.597	2.549	5.708
2	2,858.14	10.31	45,058.34	1.505	1.747	3.039
3	2,520.16	7.99	31,277.28	1.327	1.354	2.109
4	2,098.94	6.48	21,047.11	1.106	1.098	1.419
5	1,898.61	5.9	14,828.54	–	–	–

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

heterogeneity parameters found in this way and used in the model are reported in Table 3.[17]

2.3. *Elasticity of Labour*

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

[17] 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.

(*a*) *Across time* variability of aggregate hours worked is *higher* than variability of aggregate consumption.
(*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 emphasised 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 to explain 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 that in this case

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

Therefore, linear utility of leisure contradicts fact (*b*) above, because consumption is constant across agents of different wealth.

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,

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

for all j and t. This is because for high $|\gamma_l|$ agents are so averse to changes in hours worked that they are likely to choose low volatility of hours across time and they will choose 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. 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 fact (*b*) and to have a model where hours worked do not react very strongly to changes in policy. For this reason, we choose $\gamma_l = -10$ in the benchmark case which implies a very low wage elasticity of labour. This calibration is incompatible with fact (*a*).[18] As with many other parameters, we will check robustness of our main results to this choice.

[18] We check that this is the case in a model with heterogeneous agents, taxes and aggregate uncertainty in the 1995 working paper version.

2.4. *Numerical Issues*

Since before the reform the economy is at the steady state it is trivial to find the equilibrium g.

After the reform, there will be a transition period as allocations converge to the new steady state in deviations from trend. The difficulty is, therefore, finding the transition along with the labour 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 and on the model in deviations from trend are given in Appendix B. Since there is no aggregation in the model, we need to solve for the aggregate variables jointly with the individual variables. Therefore, aggregate variables are solved *jointly* with the ratios λ. In Appendix B we show that adding the ratios λ to the list of variables to be computed implies a small additional computational cost relative to a model with aggregation.

3. Results

We first show that in a homogeneous agent version of our model suppressing capital taxes causes a small improvement in welfare. This confirms the results of Lucas (1990) and Cooley and Hansen (1992) in our slightly different model and calibration. Furthermore, relative to the literature we find these gains are more robust: we find that, contrary to past results, there is 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.

3.1. *Homogeneous Agent*

3.1.1. *Replicating homogeneous agent results*

We use the benchmark parameters of Table 1. Steady state values are shown in Table 4. The first column shows values for the *status quo*, while the second column displays the values after the reform. As expected the level of capital, labour productivity and even the wage net of taxes are higher in the long run if the reform takes place. The labour tax has to increase from 23% to 37% in order to finance the capital tax cut.

Higher output in the long run does not necessarily imply that suppressing capital taxes should lead to higher welfare. 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. Therefore the transition has to be analysed explicitly.

The welfare benefits of changing the tax system are evaluated, separately for each agent. We use the standard measure given by the permanent increase in consumption that would leave each individual indifferent between the *status quo* and the reform, keeping leisure unchanged. More precisely, letting $\{c^A_{j,t}, l^A_{j,t}\}$ and $\{c^B_{j,t}, l^B_{j,t}\}$ be the equilibrium allocations before and after the reform, the welfare gain for agent j is given by π_j that satisfies

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

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%

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, slightly larger due to the higher capital taxes in our benchmark parameterisation.

3.1.2. *Emphasising 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 re-examine this result and we find that, under homogeneous agents, if the capital/output ratio is kept constant, there is a gain from suppressing capital taxes even for high risk aversion. 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 4.

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 labour to be more elastic, increasing the costs of a higher labour 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 *et al.* (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 in Table 5: even though $\gamma_c = -8$ still shows a small gain in utility due to our slightly different model and calibration, 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 the model in deviations from trend the effective discount factor becomes $\tilde{\delta} \equiv \delta\mu^{\gamma_c+1}$ (see Appendix B). 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. Table 5 shows that the steady state capital for $\gamma_c = -11$ is about one fifth of the capital for log utility. This means that for $\gamma_c = -11$ labour at the *status quo* is much less productive than in the log utility case and it

Table 5

Utility Gain from Suppressing Capital Taxes, Homogeneous Agent, Varying γ_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 labour tax rate after suppressing capital taxes in this case, while π_H measures the welfare gain when agents are homogeneous.

explains why the labour tax rate needs to be raised much more (to 70% instead of 37%) in order to compensate for suppressing capital taxes when $\gamma_c = -11$. Therefore changing $-\gamma_c$ relative to the benchmark case not only influences the elasticity of labour and the utility cost of the transition but it also increases the size of the distortion that labour has to suffer if capital taxes disappear.

In order to analyse the effects of increasing risk aversion in isolation we prefer to increase risk aversion without modifying the capital output ratio. For this purpose we change the scaling constant A in the production function to keep the same capital output ratio for different γ_c. The results are shown in Table 6. We now find that the gains from suppressing capital taxes are indeed lower for high risk aversion but the homogeneous consumer never loses utility from suppressing capital taxes, even for very high risk aversion.

In summary, the example discussed by Chari *et al.* certainly serves their purpose, namely, to show how ignoring the transition for optimal capital and labour taxes can result in an even lower utility than at the *status quo*. But suppressing capital taxes is always beneficial in terms of aggregate efficiency if the calibration is adjusted appropriately.

Table 6

Utility Gain from Suppressing Capital Taxes, Homogeneous Agent, Varying 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

Notes: See Table 5

3.2. *Heterogeneous Agents*

The main goal of this article is to study the welfare effects of eliminating capital taxes when agents are heterogeneous. Since this is a model where there is no aggregation it is not obvious that suppressing capital taxes will lead to higher aggregate output as it did in the model with homogeneous agents. However, probably because of the presence of complete markets, aggregate variables in the heterogeneous agent case 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. This can be seen in Figure 2, 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 labour supply. The last two graphs show how consumption and hours worked are very different for agents 1 and 5.

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 agents with a high wage/wealth ratio. This may offset the gains from the higher aggregate efficiency for these agents. Since we labelled $j = 5$ the agent with the highest wage/wealth ratio, a reduction in capital taxes is likely to lower the relative consumption of agent $j = 5$. Therefore, according to (12), suppressing capital taxes is likely to increase the ratios $\lambda_j = c_{j,t}/c_{5,t}$ for $j = 1,\ldots,4$.

Table 7 shows the effects of this redistribution of wealth by reporting equilibrium ratios of consumption and labour for different capital taxes, with labour taxes adjusted to maintain the same level of government spending in all cases. The first row corresponds to the *status quo* capital tax, so it simply describes the equilibrium consumption ratios $\lambda_j = c_{j,t}/c_{n,t}$ and labour 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 $-\gamma_l$ to match fact (*b*) in subsection 3.3.[19]

The last row of Table 7 corresponding to $\tau^k = 0$ shows the effects of suppressing capital taxes. We see that all groups $j = 1,\ldots,4$ 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 lowest wage/wealth ratio: while his consumption ratio increases by 70% (it goes from 3.23 before the reform to 5.56) the consumption ratio of the agent

[19] 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 this basic observation 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 results on the gains from suppressing capital taxes.

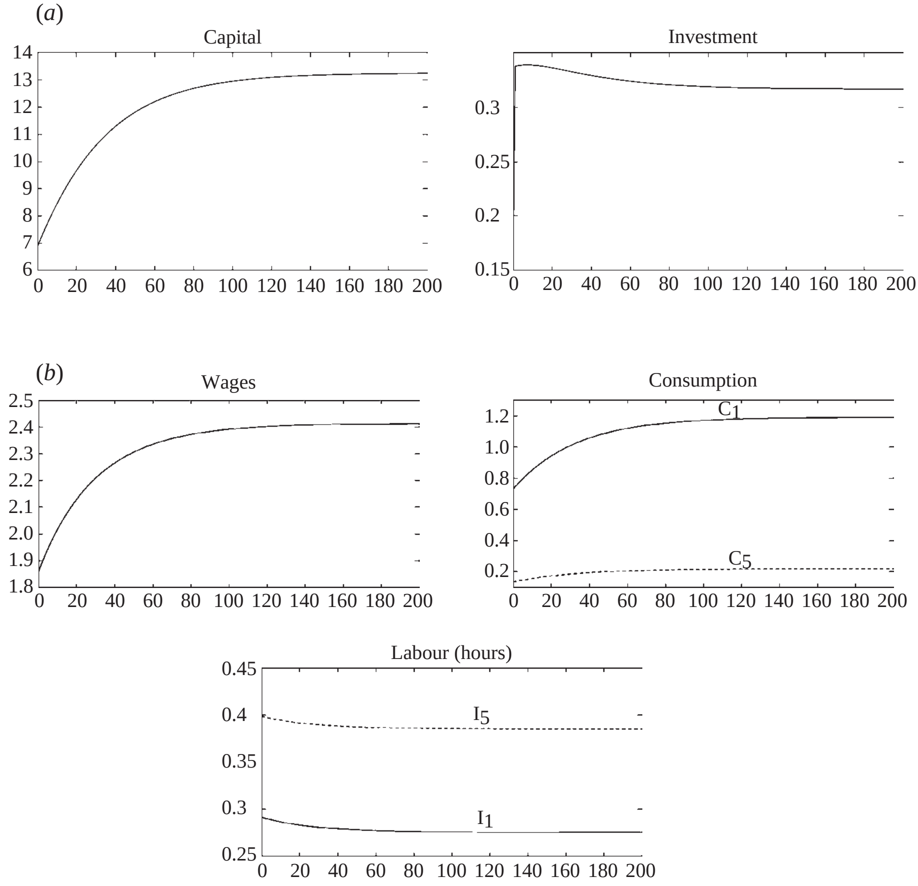

Fig. 2. *Sample Paths for* (*a*) *Capital and Investment;* (*b*) *Consumption, Wages and Hours Worked*

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 favour of the agents with a low wage/wealth ratio.

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%. We see the effect is monotone: all λs increase as capital taxes decrease. These rows will serve to understand the next Table.

It is clear from Table 7 that lowering capital taxes increases inequality. But since there is a gain in aggregate efficiency, as shown in Figure 2, it could happen that less wealthy agents experience a net gain from suppressing capital taxes. To resolve this issue we consider the 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, agents of type 5 would

Table 7
Consumption and Labour Ratios

	Wage/Wealth Partition							
New τ^k	c_1/c_5	c_2/c_5	c_3/c_5	c_4/c_5	l_1/l_5	l_2/l_5	l_3/l_5	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

experience a very large loss in welfare of 32%. 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 highest wage/wealth ratio would lose welfare, although the rest of the population would benefit.

These welfare comparisons confirm that eliminating capital income taxation at the expense of labour 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 4 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 labour 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 a sufficiently large part of the population significantly. Second, in the robustness experiments of Section 4 we will find that for slightly different parameter values the median voter often loses from suppressing capital taxes. Therefore it is not clear *ex ante* that even the median voter will favour such a reform.

4. Sensitivity Analysis

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 labour 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 for $j = 1, \ldots, 5$ show the utility gains of each agent.

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 in the literature. For each γ_c we adjust the constant A in the production function so as to keep the capital stock constant for the reasons explained in Section 3.1.2.[20]

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

Table 9

Sensitivity Analysis: Effects of Parameter Variations on Calibration and Welfare Gains of Fully Suppressing Capital Taxes

	k_{stst}	g/y	τ^l	π_H (%)	π_1 (%)	π_2 (%)	π_3 (%)	π_4 (%)	π_5 (%)
$-\gamma_c$	*Risk aversion*								
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
$-\gamma_l$	*Labour disutility*								
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
τ^k	*Status quo capital taxation*								
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

Notes: See Table 5.

[20] The following caveat is in order. While it is clear that for log utility the wage/wealth ratio is the relevant criterion for splitting the sample, with higher risk aversions this is not strictly speaking correct, since consumption may less than double when wage and wealth double. Nevertheless we maintain the calibration of heterogeneity parameters based on wage/wealth ratios. This is for three reasons: (*i*) comparability, (*ii*) simplicity, (*iii*) because this is probably a reasonable approximation to the actual equivalent agents. Probably, capturing the relevant joint distribution exactly with high risk aversion requires a more elaborate criterion than the one used in the rest of the article.

the size of welfare gains or losses is exaggerated by increasing risk aversion.[21] 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 likely values of risk aversion.[22]

It is intuitive that higher risk aversion should increase the inequality 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 welfare to gain from suppressing capital taxes. But it is also well known that the wage elasticity of labour is higher for higher risk aversion. This means that for higher $-\gamma_c$ labour goes down more steeply for a given increase in labour taxes and in order to meet the budget constraint the government needs a larger labour tax hike after the reform. As can be seen from Table 9, for a risk aversion of 1 we have $\tau^l = 0.37$ after the reform but for risk aversion of 4 we have $\tau^l = 0.46$. Agents with high wage/wealth ratio have to pay more taxes when risk aversion is higher and they lose relatively more. Also, since labour is more elastic for high risk aversion, the increase in labour taxes is more distortionary and more costly in terms of welfare for the reasons usually considered in public finance taxation.

We also consider robustness to the value of γ_l. As we explained in Section 3 the choice for the benchmark case is questionable because it would fail to account for the variability of hours worked across time in a stochastic version of the model; furthermore, it implies a wage elasticity of about 0.1 which is lower than usually estimated for the aggregate economy. We see from Table 9 that lower values of $-\gamma_l$ (and, therefore, closer to those used in the RBC literature) only exaggerate the inequality generated by suppressing capital taxes. As is well known, lower $-\gamma_l$ implies higher wage elasticity of labour and 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 (see discussion in footnote 11 for references), 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 lose (gain) less for lower status quo 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 low to begin with 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.

[21] Only the results up to $\gamma_c = -4$ are reported because the algorithm failed to converge for higher levels of risk aversion. We do not know if this is a failure of the algorithm or, more likely, this happens because there is no equilibrium with zero capital taxes, that is, there is no way to collect enough from only labour taxes in order to maintain g.

[22] For the case considered in Chari *et al.* (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 $\pi_1 = 64.65\%, \pi_2 = 22.12\%, \pi_3 = -3.15\%, \pi_4 = -22.88\%, \pi_5 = -68.49\%$.

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.

It is clear that the results are very robust. If anything, the benchmark calibration understates the redistributive effects of suppressing capital taxes.

5. Conclusion

The Chamley (1986) and Judd (1985, 1987) results say that in a model with heterogeneous agents and distortionary taxes all Pareto optimal allocations have the property that capital taxes disappear in the long run, even if the planner cares mostly about workers. One may wonder if these long-run results could be implemented immediately and if suppressing capital taxes could benefit all agents. We explore whether this is the case in a model with heterogeneous agents. Our model is as close as possible to that of Chamley (1986) and Lucas (1990) so as to explore in isolation the effects of heterogeneity.

We find that if capital taxes were suppressed and the lost revenue was compensated by higher labour 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 reasonable values of relative risk aversion, agents in the lowest quintile of the population lose 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 lose from the tax reform depends very much on the parameter values chosen for the model. We find that for reasonable levels of risk aversion the median voter would lose from the reform but for log utility it would gain. 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.

We find that there is an aggregate efficiency gain even with very high risk aversions but that, in this case, the redistributive effect is even larger.

In this sense, for the issue of capital taxation, 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 labour 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. Therefore the heterogeneity parameters in our model attempt to reproduce the features of the distribution of wage/wealth ratios.

Our intention is to examine the effect of heterogeneity in isolation, therefore we stay as close as possible to the model of Chamley throughout the article. Along the way we find 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 out that the standard neoclassical model cannot simultaneously match the observed volatility of hours worked and consumption across time and the variation of these variables 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, 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 person. A better modelling of families of different types would be crucial. Finally, the model has a difficult time explaining total wealth held by all agents and total capital income, due to the fact that all assets in our model yield very similar returns.

This indicates that there is enormous scope for future research in studying tradeoff between efficiency and equity when considering changes in the tax code with equilibrium models and heterogeneous agents.

Appendix A

Calibration of 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, labour 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 $\phi_j/k_{j,-1}$ ratios.

For the actual calibration we need to estimate the relative permanent consumption of different types of agent. For this purpose we compare the total labour and capital income of different groups and identify the ratio of income to the ratio of consumption.

PSID provides data on labour 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, CDs, IRAs 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[23].
  - Stocks, Mutual Funds: S&Ps common stock price index. (Dividends are reported as asset income in the category of 'cash assets'.)
  - Total real estate:[24] 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 cars, mobile homes etc. We do not impute any rent for this category.

Appendix B

Numerical Algorithm

We describe in detail here how we solve for the equilibrium quantities after the reform that suppresses capital taxes.

At the end of Section 1 we show the equations that characterise the sequence $\{(c_{j,t}, l_{j,t})_{j=1}^{n}, k_t\}_{t=0}^{\infty}$. To allow for a numerical solution we need to convert the model in deviations from trend, in this way a steady state can exist and we can find transitions to this steady state.

Let deviations from trend be given by $\tilde{c}_{j,t} = c_{j,t}/\mu^t, \tilde{k}_t = k_t/\mu^t, \tilde{e}_t = e_t/\mu^t$ and so on. Standard algebra shows that these satisfy

$$\tilde{c}_t + g + \tilde{k}_t - (1-\tilde{d})\tilde{k}_{t-1} = A\tilde{k}_{t-1}^{\alpha}\tilde{e}_t^{1-\alpha} \tag{14}$$

$$\tilde{c}_{n,t}^{\gamma_c} w_t (1-\tau^l)\phi_n = B(1-l_{n,t})^{\gamma_l} \tag{15}$$

$$\tilde{c}_{n,t}^{\gamma_c} = \tilde{\delta}\tilde{c}_{n,t+1}^{\gamma_c}[(\tilde{r}_{t+1} - d/\mu)(1-\tau^k) + 1/\mu] \tag{16}$$

for $\tilde{r}_t = r_t/\mu, \tilde{d} \equiv 1-(1-d)/\mu$ and $\tilde{\delta} \equiv \delta\mu^{\gamma_c+1}$.

[23] All rates of return or price series were extracted from CITIBANK.
[24] As the difference between real estate value and principal mortgage remaining.

Notice that $\tilde{d}$ does not subtitute the original depreciation rate d everywhere. In particular, in the FOC with respect to capital, we have d/μ instead.

The present value budget constraints can be rewritten in terms of deviations from trend as

$$\sum_t \tilde{\delta}^t \left(\frac{\tilde{c}_{n,t}}{\tilde{c}_{n,0}}\right)^{\gamma_c} [\tilde{c}_{j,t} - \phi_j w_t l_{j,t}(1-\tau^l)] = k_{j,-1}\mu\, [1/\mu + (\tilde{r}_0 - d/\mu)(1-\tau^k)] \text{ for } j = 1, 2, \ldots, n. \quad (17)$$

Finally, for the welfare calculations we use the equality

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

The numerical problem can be further simplified by noting that, for candidate values $\lambda_1, \ldots, \lambda_{n-1}$ we can use (9) to substitute out consumption and labour in (16) for agents $j = 1, \ldots, n-1$ in terms of $\{\tilde{c}_{n,t}, l_{n,t}\}_{t=0}^{\infty}$ and the λs.

Therefore the numerical problem at hand reduces to the following: given τ^k and g, find three sequences $\{\tilde{c}_{n,t}, l_{n,t}, \tilde{k}_t\}_{t=0}^{\infty}$, plus n constants $(\lambda_1, \ldots, \lambda_{n-1}, \tau^l)$ such that (14), (15), (16) hold for all t and (17) hold for all j.

We convert this into a finite problem by fixing large T and computing a sequence that satisfies:

(*a*) (14), (16), (15) for $t = 0, \ldots, T-1$
(*b*) (17) for $j = 1, \ldots, n$
(*c*) Variables dated $t > T-1$ are set at steady state.

Notice that (*a*) provides $3T$ equations and (*b*) provides n additional equations. We have $3T$ unknowns in $\{\tilde{c}_{n,t}, l_{n,t}, \tilde{k}_t\}_{t=0}^{T-1}$ plus n unknowns in $(\tau^l, \lambda_1, \ldots, \lambda_{n-1})$. This gives $3T + n$ unknowns and the same number of equations. We know this system of equations cannot be solved exactly, for $\tilde{k}_T$ cannot be at steady state unless the initial capital is at steady state but the system can be solved approximately by various numerical solution methods for solving non-linear systems of equations. As $T \to \infty$ we can potentially obtain an arbitrarily accurate approximation. We use $T = 200$ and check with 250 for robustness. From the graphs in Figure 2 we see that this allows the solution to reach steady state.

Notice that conditional on the model being at steady state after T periods infinite discounted sums involved in the calculations can be computed exactly.

It should be clear, therefore, that we do not use any aggregation result: aggregate capital and consumption are determined jointly with the λ's. Notice that adding heterogeneity means having to solve for $3T + n$ variables instead of $3T + 1$ in the homogeneous agent case. Therefore, despite the lack of aggregation, the increase in the computational cost from adding heterogeneity is negligible.

Universitat Pompeu Fabra and Barcelona GSE
Universitat Autònoma de Barcelona, CEPR and Barcelona GSE
Universitat Pompeu Fabra and Barcelona GSE

Submitted: 30 July 2008
Accepted: 28 May 2009

References

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

Carey, D., and Tchilinguirian, H. (2000). 'Average effective tax rates on capital, labour and consumption', OECD Economics Department Working Paper No. 258.

Chamley, C. (1986). 'Optimal taxation of capital income in general equilibrium with infinite lives', *Econometrica*, vol. 54(3), pp. 607–22.

Chari, V.V., Christiano, L.J. and Kehoe, P.J. (1994). 'Fiscal policy in a business cycle model', *Journal of Political Economy*, vol. 102, pp. 617–52.

Chari, V.V. and Kehoe P.J. (1999). 'Optimal fiscal and monetary policy', in (J. Taylor and M. Woodford, eds.), *Handbook of Macroeconomics*, pp. 1671–745, Amsterdam: North Holland.

Conesa J.C., Kitao, S. and Krueger, D. (2009). 'Taxing capital? Not a bad idea after all!', *American Economic Review*, vol. 99(1), pp. 25–48.

Cooley, T.F. and Hansen, G.D. (1992). 'Tax distortions in a neoclassical monetary economy', *Journal of Economic Theory*, vol. 58(2), pp. 290–316.

Correia, I.H. (1999) 'On the efficiency and equity trade-off', *Journal of Monetary Economics*, vol. 44 (3), pp. 581–603.

Domeij, D. and Heathcote, J. (2004). 'On the distributional effects of reducing capital taxes', *International Economic Review*, vol. 45 (2), pp. 523–54.

Feldstein, M.S., Dicks-Mireaux, L. and Poterba, J. (1983). 'The effective tax rate and the pretax rate of return', *Journal of Public Economics*, vol. 21, pp. 129–58.

Flodén, M. (2009). 'Why are capital income taxes so high?', *Macroeconomic Dynamics*, vol. 13(3), pp. 279–304.

Garcia-Milà, T., Marcet, A. and Ventura, E. (1995). 'Supply side interventions and redistribution', Working Paper Universitat Pompeu Fabra, No. 115.

Greenwood, J., Rogerson, R. and Wright, R. (1995). 'Household production in real business cycle theory', in (T.F. Cooley, ed.), *Frontiers of Business Cycle Research*, pp. 157–74, Princeton: Princeton University Press.

Hansen, G.D. (1985). 'Indivisible labour and the business cycle', *Journal of Monetary Economics*, vol. 16(3), pp. 309–28.

Joines, D.H (1981). 'Estimates of effective marginal tax rates on factor incomes', *Journal of Business*, vol. 54(2), pp. 191–226.

Jones, L.E., Manuelli, R.E. and Rossi, P.E. (1993). 'Optimal taxation in models of endogenous growth', *Journal of Political Economy*, vol. 101(3), pp. 485–517.

Judd, K.L. (1985). 'Redistributive taxation in a simple perfect foresight model', *Journal of Public Economics*, vol. 28(1), pp. 59–83.

Judd, K.L. (1987). 'The welfare cost of factor taxation in a perfect foresight model', *Journal of Political Economy*, vol. 95, pp. 675–709.

Klein P, Krusell, P. and Ríos-Rull, V. (2008). 'Time consistent fiscal policy', *Review of Economic Studies*, vol. 75(3), pp. 789–808.

Krusell, P. and Ríos-Rull, V. (1999). 'On the size of U.S. government: political economy in the neoclassical growth model', *American Economic Review*, vol. 89(5), pp. 1156–81.

Ljungqvist, L. and Sargent, T.J. (2004). *Recursive Macroeconomic Theory*, 2nd edition, Cambridge MA: MIT Press.

Lucas, R.E. Jr. (1990). 'Supply-side economics: an analytical review', *Oxford Economic Papers*, vol. 42(2), pp. 293–316.

Maliar, L. and Maliar, S. (2001). 'Heterogeneity in capital and skills in a neoclassical stochastic growth model', *Journal of Economic Dynamics and Control*, vol. 25(9), pp. 1367–97.

McGrattan, E., Rogerson, R. and Wright, R. (1997). 'An equilibrium model of the business cycle with household production and fiscal policy', *International Economic Review*, vol. 38(2), pp. 267–90.

Mendoza, E., Razin, A. and Tesar, L. (1994). 'Effective tax rates in macroeconomics. Cross country estimates of tax rates on factor incomes and consumption', *Journal of Monetary Economics*, vol. 34(3), pp. 297–323.

Rogerson, R. (1986). 'Indivisible labour, lotteries and equilibrium', *Journal of Monetary Economics*, vol. 21(1), pp. 3–16.

Rosenthal, S.S. (1988). 'A residence time model of housing markets', *Journal of Public Economics*, vol. 36, pp. 87–109.