Institute for Empirical Research in Economics
University of Zurich

Working Paper Series
ISSN 1424-0459

Working Paper No. 354

Optimal Ramsey Tax Cycles

Marcus Hagedorn

December 2007

Optimal Ramsey Tax Cycles*

Marcus Hagedorn
University of Zurich†

October, 2007

Abstract

This paper asks whether tax cycles can represent the optimal policy in a model without any extrinsic uncertainty.

I show, in an economy without capital and where labor is the only choice variable (a Lucas-Stokey economy), that a large class of preferences exists, where cycles are optimal, as well as a large class where they are not.

The larger government expenditures are, the larger the class of preferences for which cycles are optimal becomes.

Tax cycles are also more likely to be optimal if frictions (deviations of the model from Walrasian markets) are added. While this cannot be shown in general and will not be true for arbitrary frictions, I demonstrate this in two specific worlds. I consider an economy with search frictions in the labor market, and one with frictions in the goods and credit market. A reasonable parametrization of both economies shows that results change considerably. Even with constant relative risk aversion, cycles can be optimal, whereas this class of preferences rules out cycles in the Lucas-Stokey economy.

Finally, I characterize the optimal policy. No more than two tax rates are needed to implement the Ramsey policy both in the Lucas-Stokey economy and in the model with frictions.

JEL Classification: H21, E32, E62, E63

KEYWORDS: Optimal Taxation, Tax Cycles, First-order Approach.

*I would like to thank Dirk Krüger, Stefan Niemann and the seminar participants at Penn for their helpful comments and suggestions that have been incorporated throughout the paper. Financial support from NCCR-FINRISK and the Research Priority Program on Finance and Financial Markets of the University of Zurich is gratefully acknowledged

†Institut for Empirical Research (IEW), University of Zurich, Blümlisalpstrasse 10, Ch8006 Zürich, Switzerland. Email: hagedorn@iew.uzh.ch.

1 Introduction

This paper asks whether tax cycles represent the optimal Ramsey tax policy in the simplest framework possible (Lucas and Stokey (1983), LS hereafter). There is no uncertainty, no capital or money, and households' only endogenous choices are how much labor to supply and how many bonds to buy. The government issues bonds and taxes labor income to maximize welfare, and can commit to future policies. The question is then: is it optimal, in a model without any extrinsic uncertainty, to smooth taxes as suggested by Barro (1979), or are tax cycles welfare-improving?

The explanation why cycles could represent the optimal choice is simple. Mathematically, the reason is the potential non-convexity of the Ramsey optimization problem. First-order conditions are then not sufficient to characterize the optimum: the first-order approach is invalid. The Ramsey planner takes into account the fact that a change in tax rates changes the price of consumption and thus interest rates. A cycle is optimal whenever this change in interest rates relaxes the government's budget sufficiently to compensate for the welfare loss from the induced variation in consumption, hours and wages. In particular, cycles are fully policy-induced and would not exist without policy.[1]

I show that in the LS economy, two classes of preferences exist, with different implications. One class implies that tax smoothing is optimal, whereas within the other class, tax cycles improve welfare. With separability between consumption and leisure, these two classes can be fully and sharply characterized.

(Weakly) Increasing relative risk aversion (IRRA) in consumption[2] implies the optimality of tax smoothing, whereas sufficiently strong decreasing relative risk aver-

[1] Although the question here is whether it is optimal for the Ramsey planner to implement cycles in an acyclical economy, cycles can be the market outcome in other (multi-sector growth) models (McKenzie (1986), Boldrin and Montrucchio (1995)).

[2] What matters here is the elasticity of intertemporal substitution, although the preferences (DRRA, CRRA, IRRA) are classified according to their risk aversion.

sion (DRRA), empirically the more relevant case[3], implies the optimality of a cycle.

For non-separable preferences, if consumption and leisure are substitutes, the Ramsey policy is characterized by cycles for a larger set of parameter values.

I then derive a simple criterion in subsection 3.2 to check whether a two-period cycle is welfare-improving in the LS economy. This criterion is silent on the shape of the optimal policy. Subsection 3.5 partially fills this gap. An optimal policy (after the initial period) can be implemented by no more than two different tax rates. The optimal policy is either constant or discontinuous. In particular, the optimal policy is not necessarily differentiable as Chamley (1986) assumes in order to verify that cycles are not optimal.

A natural question that then arises is whether IRRA always, in models other than that of LS, rules out cycles as part of the optimal policy. To show that the results apply in models that are more policy-relevant than LS, I show that they also apply in a model with labor market frictions (as in Pissarides (1985, 2000)), and again in Lagos and Wright (2005), a model with frictions in the goods and credit market. The advantage of choosing these models is that they are close to an LS economy in many respects, permitting analytical results to be derived, which are useful for interpreting my own findings. At the same time, exactly the same frictions which characterize these simple models are also present in much richer models used for policy analysis, so that the possibility of tax cycles being optimal cannot be ruled out *a priori* for these models either.

Labor market frictions are modeled, following Pissarides (1985, 2000), as search frictions in a search and matching model. The difference from LS is that the only (interesting) decision in LS is how many hours to work, whereas the decision in a search economy is how many vacancies to post. The theoretical results derived for LS are thus applicable. The simple economy differs from Pissarides (1985, 2000) in

[3]Vissing-Jorgenson (2002) provides evidence that richer households have a higher elasticity of intertemporal substitution than poorer ones. For this type of preferences, this implies that risk aversion decreases with wealth. See Gollier (2001) for a further discussion of DRRA and references to the literature.

terms of bargaining and in the explicit consideration of intertemporal substitution. When the model is calibrated in a standard way, I find that cycles become optimal for constant relative risk aversion (CRRA).

To model product market frictions, I consider the Lagos and Wright (2005) model, which provides a tractable way to deal with two separate frictions. The terms of trade of some transactions are determined through bargaining, and credit cannot be used to pay for these goods. In a parameterized example I find that cycles can improve welfare for *all* CRRA preferences.

Thus the belief that restricting oneself to IRRA or even CRRA could rule out cycles and potentially validate the first-order approach is not warranted. Cycles can improve welfare even in a simple model, one not too different from the LS world. Cycles are optimal for two reasons. First, frictions imply that the economy is inefficient (the first welfare theorem does not hold even without a government). If this inefficiency is strong enough, cycles can represent the optimal policy, a result that is already true in the LS economy.[4] Second, prices or wages are not formed in a perfectly competitive market; that is, marginal rates of transformation do not equal marginal rates of substitution. This deviation from LS can make prices and wages and thus tax revenue a more convex function of the allocation.[5]

My paper is related to Hassler, Krusell, Storesletten, and Zilibotti (2004) (HKSZ), who show that without geometric depreciation of capital, optimal capital income tax rates oscillate. However, their paper differs from mine in that they are not only silent on the properties of labor taxes, but also provide a quite different reason for oscillations (moreover, in HKSZ oscillations generically die out).[6] The

---

[4]In an LS economy, cycles can become the optimal policy if the allocation is sufficiently inefficient (because of higher government expenditures and thus higher taxes).

[5]A more sophisticated search economy is considered by Shimer and Smith (2001), who show that the optimal matching policy can be non stationary. However, the reason for non stationarity is quite different. Ex ante heterogeneity of workers in their model leads to a matching externality which is absent from my model.

[6]As a special case, if capital does not depreciate in the first period and fully depreciates in the second period, oscillations do not die out. In all other cases, the tax rate converges to a constant

timing of taxes is an optimal response to the different depreciation rates of old and new capital, since it allows the inelastically supplied capital stock in period 0 to be partially confiscated (although tax rates are zero in period 0).[7]

The remainder of this paper is organized as follows. Section 2 introduces the model. Section 3 considers the optimal policy and its characteristics in the basic Lucas-Stokey economy. Section 4 analyzes a search and matching model of the labor market, while Section 5 investigates the Lagos and Wright (2005) model. Section 6 concludes by summarizing the results. All proofs are delegated to the appendix.

2 The Model

In this section I describe the Lucas and Stokey (1983) economy without uncertainty. All prices are determined in competitive Walrasian markets, and households decide how many hours to work, how much to consume and how many bonds to buy. The government finances an exogenous stream of expenditures by levying distortionary labor income taxes and issuing bonds.

2.1 Private Sector

The economy is populated by a large number of identical households. Each household's preferences over a stream of consumption and leisure are described by a utility function

$$\sum_{t=0}^{\infty} \beta^t u(c_t, h_t) \tag{1}$$

level, whereby oscillations die out.

[7]Kocherlakota (2006) criticizes the arbitrary restriction of tax instruments in the Ramsey approach and argues that the only constraints in a taxation problem should be informational and/or enforcement frictions. This critique applies here as well. But, interestingly, it has been demonstrated in this literature as well that the first-order approach is not valid (Kocherlakota (2004) and Coles (2006)).

where c_t denotes consumption, h_t denotes labor and $\beta \in (0,1)$ denotes the subjective discount factor. The single-period utility function u is assumed to be strictly increasing in c_t and strictly decreasing in h_t, strictly concave and three times continuously differentiable. $u_c(t)$ and $u_h(t)$ denote the derivative of period t utility with respect to c and h respectively. There is no uncertainty. In each period $t \geq 0$, households can buy one-period government bonds B_{t+1}, which pay out B_{t+1} consumption goods in period $t+1$. The product market is perfectly competitive, and the production function is assumed to be a linear function of labor h_t, $F(h_t) = h_t$. Firms take the real wage w_t, which equals 1 in equilibrium, as given. They hire labor h_t to maximize profits. Households are paid a net wage rate $w_t \cdot (1-\tau_t)$, where τ_t is the labor tax rate at period t.

The household's flow budget constraint in period $t \geq 1$ is given by:

$$c_t + B_{t+1}/R_{t+1} \quad \leq \quad B_t + (1-\tau_t) w_t h_t, \tag{2}$$

where $1/R_{t+1}$ denotes the period-t price of a claim to one unit of consumption in period $t+1$. The left-hand side of the budget constraint represents the uses of wealth: consumption spending and bond purchases. The right-hand side shows the sources of wealth: bonds acquired in the previous period, plus labor income.

The household starts with initial wealth B_0 and is subject to the following borrowing constraint that prevents it from engaging in Ponzi schemes:

$$\lim_{t \to \infty} q_t B_t \quad \geq \quad 0, \tag{3}$$

where $q_t := \prod_{i=1}^{t} \frac{1}{R_i}$ is the period-zero price of one unit of consumption to be delivered in period t. q_0 is normalized to one.

The household chooses $\{c_t, h_t, B_{t+1}\}_{t=0}^{\infty}$ to maximize (1), subject to (2) and (3), taking as given $\{w_t, R_{t+1}, \tau_t\}_{t=0}^{\infty}$ and initial wealth B_0.

2.2 Government

The government faces a stream of unproductive public consumption $g_t = g$. This is financed by levying labor income taxes at a rate τ_t and by issuing one-period bonds. The government's sequential budget constraint is given by

$$B_{t+1}/R_{t+1} = B_t + g_t - \tau_t w_t h_t. \tag{4}$$

2.3 Equilibrium

Definition 1. *Given an initial condition B_0, a competitive equilibrium is a sequence $\{c_t, g_t = g, h_t, \tau_t, w_t, B_{t+1}, R_{t+1}\}_{t=0}^{\infty}$ such that:*

i) The government budget constraint is satisfied for all $t \geq 0$.

ii) The household's choice problem is solved.

iii) The goods market clears: $c_t + g = h_t$.

iv) Firms maximize profits.

A competitive equilibrium $\{c_t, g_t = g, h_t, \tau_t, w_t, B_{t+1}, R_{t+1}\}_{t=0}^{\infty}$ then satisfies the following conditions:

$$\begin{aligned}
c_t + g &= h_t && (5)\\
\frac{-u_h(c_t, h_t)}{u_c(c_t, h_t)} &= w_t(1-\tau_t) = (1-\tau_t) && (6)\\
c_t + B_{t+1}/R_{t+1} &= B_t + (1-\tau_t)h_t && (7)\\
u_c(c_t, h_t) &= \beta R_{t+1} u_c(c_{t+1}, h_{t+1}) && (8)\\
\lim_{t\to\infty} q_t B_t &= 0, && (9)
\end{aligned}$$

where (5) is the resource constraint, (6) describes an equilibrium in the labor market, (7) is the household's budget constraint, (8) is the consumption Euler equation, and (9) is the no-Ponzi condition.

3 Optimal Dynamic Policy

In this section I consider whether implementing a constant tax rate is an optimal strategy for the government. I derive a condition to check whether a cycle is superior, and prove the properties of such a non-deterministic policy.

3.1 The Ramsey problem

The Ramsey problem is the choice of an implementable allocation which maximizes welfare. The next two definitions state this more precisely.

Definition 2. *An allocation* $\{c_t, h_t\}_{t=0}^{\infty}$ *is implementable if* $\{\tau_t, w_t, B_{t+1}, R_{t+1}\}_{t=0}^{\infty}$ *exists, such that* $\{c_t, g_t = g, h_t, \tau_t, w_t, B_{t+1}, R_{t+1}\}_{t=0}^{\infty}$ *is a competitive equilibrium.*

Definition 3. *A Ramsey solution is an allocation that maximizes welfare over all implementable allocations.*

I now reproduce the well-known primal-form representation as described in Lucas and Stokey (1983). The basic idea is to use first-order conditions to eliminate all prices and taxes from the equilibrium conditions. Only the two variables consumption and hours appear in the resulting primal form. This method allows all implementable allocations to be characterized by only two equations. The first of these is the resource constraint (5). To derive the second equation, I start with the intertemporal budget constraint of the representative household:

$$\sum_{t=0}^{\infty} q_t(c_t - (1-\tau_t)h_t) = B_0.$$

Since the individual's consumption Euler equation implies that

$$q_t = \beta^t \frac{u_c(c_t, h_t)}{u_c(c_0, h_0)},$$

and as the intratemporal condition implies

$$-\frac{u_h(c_t, h_t)}{u_c(c_t, h_t)} = 1 - \tau_t,$$

I can substitute for q_t and τ_t. The budget constraint can then be written as

$$\sum_{t=0}^{\infty} \beta^t \frac{u_c(c_t, h_t)}{u_c(c_0, h_0)} \left[c_t + \frac{u_h(c_t, h_t)}{u_c(c_t, h_t)} h_t \right] = B_0,$$

which is equivalent to the implementability constraint

$$u_c(0) B_0 = \sum_{t \geq 0} \beta^t (u_c(t) \cdot c_t + u_h(t) \cdot h_t).$$

The following proposition presents the primal-form characterization for this economy, which says that the implementability constraint and the resource constraint are necessary and sufficient for implementability.

Proposition 1. *For any B_0, an allocation $\{c_t, h_t\}_{t=0}^{\infty}$ is implementable if and only if*

$$u_c(0) B_0 = \sum_{t \geq 0} \beta^t (u_c(t) \cdot c_t + u_h(t) \cdot h_t) \tag{10}$$

$$c_t + g = h_t. \tag{11}$$

Proof: See Lucas and Stokey (1983) or Chari and Kehoe (1999).

The *Ramsey problem* can be stated as choosing consumption c_t and labor h_t to maximize (1), subject to equations (10) and (11).[8] I restrict myself to $B_0 = 0$ in what follows since this makes period $t = 0$ look like all other periods ($t > 0$). All theoretical results hold without this assumption, and the quantitative implications will be discussed later.

3.2 A Criterion for Optimal Cycles

The Ramsey problem is stationary, and a reasonable guess is that a time-invariant allocation is optimal.

[8] An upper-bound on hours ensures that the objective function is bounded, and I proceed under the assumption that g is not too large, so that the set of feasible policies is not empty. The same mathematical theorems (for example Tychonoff's theorem) as applied in Aiyagari (1994) then imply existence.

Definition 4. *A time invariant allocation* (c, h) *is an implementable sequence* $\{c_t = c, h_t = h\}_{t=0}^{\infty}$.
An optimal time invariant allocation (h^{TI}, c^{TI}) *achieves the highest welfare among all time invariant policies.*

Optimality in the definition is conditional on being time-invariant (TI). The logic of the Laffer curve implies that there are multiple TI allocations.[9] The optimal TI allocation picks the solution with the highest welfare. h^{TI} then (not uniquely) solves $g = h \cdot (1 + u_h(h - g, h)/u_c(h - g, h))$, and c^{TI} equals $h^{TI} - g$.
This section's theorem will show when a TI allocation and thus a TI policy is not the solution to the Ramsey problem. Labor supply and consumption are not constant in this case.

What does a simple time-varying policy look like? The main features of such a policy can be conveyed in a two-period model. In the first period, labor taxes are lowered and thus tax revenue decreases, so that debt has to be issued to pay for government expenditure. In the second period, the labor tax rate is increased to repay the debt (plus interest payments). The size of the necessary tax increase in the second period depends on the level of the real interest rate. To understand how this policy affects the interest rate, note that the labor supply and thus consumption increase in the first period and decrease in the second. The consumption Euler equation then implies that the interest rate decreases.

This reasoning suggests that two numbers are crucial if cycles are to be optimal. The first of these number is the elasticity of intertemporal substitution, which determines how much interest rates drop in response to the tax changes described in the previous paragraph. The second number is the wage elasticity of hours worked. This determines how much labor supply decreases (increases) and tax revenues increases (decreases) if the tax rate is increased (decreased).

To assess the (non)optimality of a TI allocation, two concepts have to be intro-

[9]This multiplicity does not imply the existence of a cycle.

duced. Let

$$CS(h_t) \quad = \quad u_c(h_t - g, h_t) \cdot (h_t - g) + u_h(h_t - g, h_t) \cdot h_t$$

be the current value of period t surplus[10] and

$$CU(h_t) = u(h_t - g, h_t)$$

be the current value of period t utility.

The convexity of $CS(\cdot)$ is a necessary (but not sufficient) condition for a TI policy not to be optimal, since otherwise first-order conditions would be sufficient to characterize the Ramsey solution. With convexity of CS, there can be more than one solution to the first-order conditions.

A cyclical policy increases welfare only if the average labor supply increases over the cycle. The increase in welfare is greater the higher $\mathbf{CU}'$ is at the TI solution h^{TI}.[11] And CU' is higher the further h^{TI} is away from the efficient solution. But utility is concave, and thus households prefer a constant consumption level to a cyclical consumption level if both choices have the same average value. This aversion to cycles increases the higher $\mathbf{CU}''$ is.

To implement a higher labor supply level, taxes have to be lowered. Since average labor supply increases, labor taxes therefore have to decrease on average. The size of this decrease depends on $\mathbf{CS}'$. The convexity of CS has to ensure that the present value budget is nevertheless balanced. The higher $\mathbf{CS}''$ is, the greater is the increase in the government's tax revenues.

To summarize, a tax cycle is optimal if four conditions are met: the time-invariant policy is sufficiently inefficient (CU' high); the utility function is not too concave ($-CU''$ low) so that the aversion to cycles is not that large; the labor supply is sufficiently responsive to tax changes ($-CS'$ low); and tax revenues are sufficiently convex (CS'' high). Since these four variables (CU', CU'', CS', CS'') all depend on

[10]Note that CS is indeed the government's primary surplus since (6) implies that $CS(h) = u_c(h-g,h)(h-g-(1-\tau)h) = u_c(h-g,h)(\tau h - g)$.

[11]$f'(f'')$ denotes the first (second) derivative of a one-dimensional function f.

the utility function, they cannot be changed independently from each other. For example, for a utility function of the form $\frac{c^{1-\sigma}}{1-\sigma} - \chi(h)$, a lower elasticity of substitution $\frac{1}{\sigma}$ increases both $-CU''$ and CS''. However, two variables, CS'' and CU', turn out to be key. CS'' has to be positive to render the problem non-convex. In addition, the economy has to be sufficiently inefficient (CU' high), either due to high government expenditures or due to other "frictions", as for example in Sections 4 and 5.

The following theorem presents a condition which indicates when a cycle is welfare-improving. This criterion summarizes the previous arguments and is analyzed in the remainder of this subsection.

Theorem 1. *Let (h^{TI}, c^{TI}) be the optimal TI allocation.*

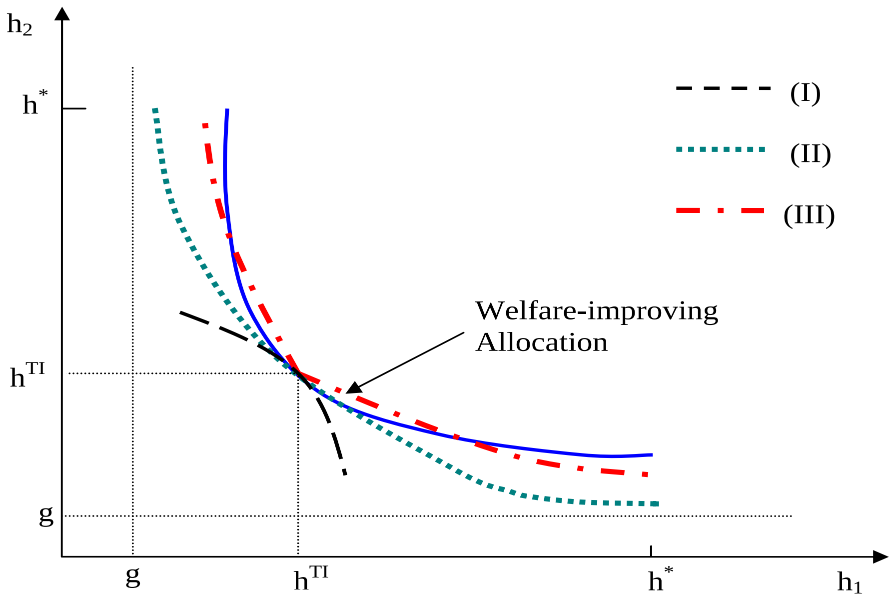
$$\frac{CS''}{CU''}(h^{TI}) - \frac{CS'}{CU'}(h^{TI}) < 0 \qquad (*)$$

implies that the optimal TI policy is not the solution to the Ramsey problem. A welfare-improving two-period cycle exists.

Derivatives are taken with respect to h. The criterion in the theorem is equivalent to a second-order condition of the constrained maximization problem. The proof uses a perturbation argument to show that a two-period cycle is welfare-improving. Condition $(*)$ can easily be rewritten in terms of primitives. CU', CU'', CS', CS'' depend on the utility function and h^{TI} is a solution to a first-order condition that purely depends on primitives.

Figure 1 graphically illustrates the possibility of a cycle in a two-period model. Labor supply in the first period h_1 is on the x-axis, and labor supply in the second period h_2 is on the y-axis. The efficient level of labor supply is h^*, and h^{TI} is the TI policy. The solid line is the indifference curve of the representative household through the point (h^{TI}, h^{TI}). Allocations with higher labor supply h generate higher utility as long as h is smaller than h^*. The graph also shows curves in the (h_1, h_2) space which generate the same tax revenue as (h^{TI}, h^{TI}). There are three possibilities for the shape of this curve. In case (I) the curve is concave. In case (II) the curve is convex but "less convex" than the indifference curve. In case (III) the curve is

Figure 1: Optimal taxation with two periods



convex and "more convex" than the indifference curve at (h^{TI}, h^{TI}). Eventually the tax curve has to intersect the indifference curve when h becomes too small and the indifference curve is again more convex. But what matters for condition $(*)$ is the curvature at h^{TI}.

In the first two cases a TI policy is optimal. There is no combination of h_1 and h_2 that generates the same tax revenue as (h^{TI}, h^{TI}) and higher utility. In case (III), such a combination exists, as indicated by the arrow. A two-period cycle then improves welfare. This is true whenever the tax curve is more convex than the indifference curve. Mathematically both curves describe h_2 as a function of h_1. The elasticity of this function equals $\frac{-CU''h^{TI}}{CU'}\frac{(1+1/\beta)}{\beta}$ for the indifference curve and $\frac{-CS''h^{TI}}{CS'}\frac{(1+1/\beta)}{\beta}$ for the tax curve. The graphical condition - namely that the tax curve is more convex than the indifference curve - can be expressed as $\frac{-CS''h^{TI}}{CS'} > \frac{-CU''h^{TI}}{CU'}$ and is thus equivalent to condition $(*)$.

A welfare-improving allocation is located above the indifference curve but below the tax curve. However, condition $(*)$ is only a local test since it compares the elasticities of CU and CS around h^{TI}. In particular, condition $(*)$ does not hold globally, so that a corner solution does not improve welfare.

3.3 Time-invariant Ramsey Policy

There is a class of utility functions such that the optimal TI policy is the Ramsey policy. To my knowledge this result is new. It is not trivial since the implementability constraint need not be convex (for example as in Case (II) in Figure 1). Furthermore, a result that non-steady-state capital income tax rates are zero (Chari and Kehoe (1999)) does not automatically imply that labor tax rates are constant. The well-known results by Judd (1985) and Chamley (1986), namely that steady-state capital income tax rates are zero, does not rule out cycles either. In case the optimality criterion implies that a (discontinuous) cycle should be implemented, a constant allocation is suboptimal. Thus a restriction to steady states is restrictive and cannot be detected through computing dynamics (eigenvalues) around this steady state because of the discontinuity.

Theorem 2. *Let* $u(c_t, h_t) = w(c_t) - \chi(h_t)$ *where*

- *w exhibits (weakly) increasing relative risk aversion (IRRA)* **and**
- *$\chi_h(h_t) \cdot h_t$ is a convex function of h_t.*

Then the optimal TI policy is the Ramsey policy.

This theorem applies to classes of preferences considered by the literature (e.g. Lucas and Stokey (1983) and Chari and Kehoe (1999)). The next two examples provide several functional forms for w and χ which satisfy the assumptions of theorem 2.

Example (Functional forms for w in theorem 2)**.**

i) $w(c_t) = (c_t^{(1-\sigma)} - 1)/(1-\sigma)$ *(CRRA)*

ii) $w(c_t) = -exp(-\sigma \cdot c_t)$ *(CARA),*

iii) Quadratic w.

The problem is not necessarily convex if $\chi_h(h_t) \cdot h_t$ is not convex. Although this assumption does not follow from standard assumptions on primitives, examples of functional forms for χ are easily found.

Example (Functional forms for χ in theorem 2)**.**

i) $\chi(h_t) = \alpha \cdot h_t$*, for some* $\alpha > 0$*.*

ii) $\chi(h_t) = h_t^\alpha$*, for some* $\alpha > 1$*.*

iii) $\chi(h_t) = -(T-h)^{1-\theta}/(1-\theta)$*, for some* $\theta > 0$ *and a time endowment* T*.*

3.4 Time-varying Ramsey Policy

The last section showed that for separable preferences which exhibit IRRA, time variation is not optimal. However, this conclusion does not necessarily hold true for the presumably relevant case of DRRA.[12] The next theorem shows that, for the Hansen (1985) preferences, sufficiently strong DRRA leads to the non-optimality of a TI policy. For $u(c_t, h_t) = w(c_t) - A \cdot h_t$ let $\kappa_c = \frac{w''' \cdot c}{w''}$, $\eta_c = \frac{w'' \cdot c}{w'}$ and τ^{TI} be the tax rate in the optimal TI policy. Set κ and η to zero if the denominator equals zero.

Theorem 3. *Let* $u(c_t, h_t) = w(c_t) - A \cdot h_t$ *and* κ_c, η_c *and* τ^{TI} *be as defined before. Then*

$$1 + \kappa_c < \eta_c / \tau^{TI} \qquad (**) \tag{12}$$

implies that the optimal TI policy is not the solution to the Ramsey problem. A welfare-improving two-period cycle exists.

DRRA is equivalent to $1 + \kappa_c \leq \eta_c$. The criterion says that DRRA has to be sufficiently strong. How strong depends on the size of τ^{TI}. If τ^{TI} approaches 1, a cycle would be optimal for all DRRA preferences, whereas $\tau^{TI} = 0$ implies efficiency and no cycle at all. A cycle is welfare-enhancing if a variation in consumption leads to a variation in prices that is high enough to compensate for the household's aversion

[12]See for example Gollier (2001) and Vissing-Jorgenson (2002), who argue that decreasing relative risk aversion seems to be the empirically relevant case. Note that, for the preferences considered here, the elasticity of intertemporal substitution is the inverse of the coefficient of relative risk aversion.

to this variation in consumption over time. Since the period t price of consumption equals w', the size of the price variation is high if w''' is high. This is exactly what DRRA achieves.

The previous theorem only considers linear disutility of labor. The condition provided there is a special case of the criterion in the next theorem, which makes no assumptions regarding the functional form of the disutility of labor. It still states that sufficiently strong DRRA leads to cycles. How strong now also depends on the curvature of the disutility of labor. As a result, the condition becomes more complicated than condition $(**)$, which purely focused on the role of DRRA. For $u(c_t, h_t) = w(c_t) - \chi(h_t)$, let τ^{TI} be the tax rate in the optimal TI policy and define $\kappa_c = \frac{w'''\cdot c}{w''}$, $\eta_c = \frac{w''\cdot c}{w'}$, $\kappa_h = \frac{\chi'''\cdot h}{\chi''}$ and $\eta_h = \frac{\chi''\cdot h}{\chi'}$ if the denominator is not zero. If the denominator equals zero, set κ or η to zero.

Theorem 4. *Let $u(c_t, h_t) = w(c_t) - \chi(h_t)$ and and $\kappa_c, \eta_c, \kappa_h, \eta_h$ and τ^{TI} be as defined before. Then*

$$1 + \frac{\kappa_c\eta_c - \kappa_h\eta_h(1-\tau^{TI})^2}{\eta_c - \eta_h(1-\tau^{TI})^2} + \frac{\eta_h(1-\tau^{TI}) - \eta_c}{\tau^{TI}} < 0 \qquad (***) \tag{13}$$

implies that the optimal TI policy is not the solution to the Ramsey problem. A welfare improving two-period cycle exists.

Note that utility functions that meet the assumptions of theorem 2 (and thus imply that a TI policy is optimal) do not satisfy the criterion in theorem 4. The condition that $\chi_h(h_t) \cdot h_t$ is convex is equivalent to $\kappa_h \geq -2$. The next proposition establishes this claim and confirms the intuition that linear utility makes cycles more attractive, as long as $\kappa_h \geq -2$ or, equivalently, as long as $\chi_h(h_t) \cdot h_t$ is convex.

Proposition 2. *Under the assumptions of theorem 4, it holds that:*

- *If $\kappa_h \geq -2$ and $1 + \kappa_c - \eta_c \geq 0$ (IRRA), then condition $(***)$ is not satisfied.*

- *If $\kappa_h \geq -2$ and if condition $(***)$ holds, then condition $(**)$ holds.*

The second part of the theorem states that, if the criterion in theorem 4 indicates that a cycle improves welfare, then a cycle also improves welfare if disutility is linear (for the same tax rate τ^{TI} and the same κ_c and η_c).

Abandoning the assumption of separability between labor and consumption in the utility function can also render a TI policy nonoptimal. This is for example the case when $u(c,h) = (c \cdot (3-h)^2)^{(1-\sigma)}/(1-\sigma)$. For this specification of u, cycles are more attractive for a less efficient allocation (due to a higher g), as predicted by criterion ($*$). But for realistic values of government expenditures g, cycles require values of $\sigma > 10$, which is within the range provided by Hall (1988), but higher than the values usually used in calibration exercises.

3.5 Properties of the Ramsey Policy

Theorem 1 shows the nonoptimality of a TI policy through the existence of an improving cycle of length two. It is silent about the optimality of this cycle. The following theorem partially fills this gap. It states that the optimal policy implements only two different tax and labor supply levels.

Theorem 5. *The welfare of the optimal Ramsey solution can be arbitrarily well approximated by a policy that implements only two different levels of h if $\beta > 1/2$.*[13]

The idea of the proof is to allow for a specific kind of randomization to convexify the problem and turn it into a linear programming problem. A well-known result in linear programming shows that the support of the optimal (randomizing) policy is not larger than two. It then remains to be shown that the optimal probability distribution of labor supply levels can be implemented by a sequence of tax rates. It is this last step where, because of the specific choice of randomization, the assumption $\beta > 1/2$ becomes relevant. An approximation is needed to rule out an infinite number of different tax rates.

So far I have characterized two classes of preferences, one for which time invariance is optimal, and one for which it is not. A possible response could be to

[13] Arbitrary approximation means that the difference between the welfare level (and the allocations) of the optimal Ramsey policy and the welfare level of a policy with two tax rates can be made smaller than any $\epsilon > 0$.

restrict oneself to the class of preferences where first-order conditions are sufficient, and ignore the problem from then on. This may be a reasonable approach if the model being studied is the LS economy, yet this is very often not the case. Various frictions are built into models in monetary economics (see for example Woodford (2003) and Lagos and Wright (2005)) and many labor market models incorporate search frictions (Pissarides (2000)). As the next two sections show, the hope that results from the LS economy carry over to these more elaborated models, which are characterized by several frictions, is not warranted.

I consider two specific economies to stress the role of frictions. I first choose a simple version of a (labor) search and matching economy, in order to approximate the LS economy as close as possible. In both models there is only one interesting market, the labor market, either competitive or characterized through search frictions. Except for this difference in modeling the labor market, the models are identical. In particular, there is only one relevant decision variable every period: either the number of hours worked or the number of vacancies posted. This makes theorems 1 and 5 applicable for the search and matching economy and more generally makes a comparison with previous results possible.

I then consider optimal cycles in the Lagos and Wright (2005) model. I choose this model since it provides an analytically tractable framework to deal with frictions in the product and credit market rather than frictions in the labor market. The terms of trade of some goods are determined through bargaining, and credit cannot be used to pay for them. Again this economy is simple enough to make theorems 1 and 5 applicable. In addition, the fact that LS arises as a limiting case of Lagos and Wright (2005) (when frictions vanish) allows me to answer the question how the size of frictions affects the likelihood of an optimal tax cycle.

For both types of frictions - in the labor and in the goods market - I find that a reasonable parametrization leads to the conclusion that cycles improve welfare. This suggests that tax cycles are optimal in a large class of models with labor or product market frictions.

4 A Search and Matching Economy

In this section I consider a search and matching model with the following features. The equilibrium is inefficient even without government expenditures. As in the LS world, it is an infinite replication of a one-shot economy. Aggregate hours change only because individuals leave and enter employment (as in Hansen (1985) and Rogerson (1988)).

In the chosen environment, there is a continuum of infinitely lived workers and a continuum of infinitely lived firms, each of measure one. At the beginning of each period they are, for the sake of simplicity, all unmatched. This eliminates unemployment as a state variable and makes the economy a replication of one-shot economies. The production function remains unchanged, and workers derive utility $u(c_t)$ from consuming c_t at time t. The discount factor is still denoted by β. Workers can be matched or unmatched. If matched they produce one unit of output, are paid a wage w_t and receive a net wage $w_t \cdot (1 - \tau_t)$. Matched and unmatched households have the same level of consumption every period. There is a market that insures households against the possibility of not meeting an employer. Each household which meets a firm pays a transfer to those which do not meet a firm such that the consumption level is equalized. Unmatched workers do not produce and purely consume the transfer from this insurance.

Firms make profit $1 - w_t$ per worker and post vacancies at a cost of k. Free entry implies that the expected value of an open vacancy is zero.

The number of matches at t is given by a constant returns to scale Cobb-Douglas matching function $m(1, v_t) = \chi \cdot 1^{\alpha} v_t^{(1-\alpha)} = \chi \cdot v_t^{(1-\alpha)}$. Employment E_t equals $m(1, v_t)$. The probability that a worker will be matched in period t equals $E_t = m(1, v_t)$, and the probability that a vacancy will be filled equals $m(1, v_t)/v_t = \chi \cdot v_t^{-\alpha}$. I assume that $m(1, v) = \chi \cdot v_t^{1-\alpha} \leq \min(v, 1)$. Wages are formed through bargaining in an alternating offers game as in Binmore, Rubinstein, and Wolinsky (1986). Hall and Milgrom (2006) show that this bargaining game results in a wage $\frac{1}{2}(1 + \frac{b}{1-\tau_t})$, where

b is the worker's utility from perpetual disagreement.[14] The after-tax wage I_t then equals $\frac{1}{2}(1-\tau_t+b)$ and, because of income pooling, each household's after-tax income equals $I_t E_t$.

The government's policy options and the household budget constraint remain unchanged.

Definition 5. *An equilibrium is an initial condition* B_0 *and a sequence* $\{c_t, g_t = g, v_t, E_t, \tau_t, B_{t+1}, R_{t+1}, I_t, w_t\}_{t=0}^{\infty}$*, such that:*

i) The government budget constraint is satisfied:
$B_{t+1}/R_{t+1} = B_t + g_t - \tau_t \frac{I_t}{(1-\tau_t)} E_t$.

ii) The resource constraint holds: $c_t + g + kv_t = E_t$.

iii) The household budget constraint is satisfied:
$c_t + B_{t+1}/R_{t+1} = B_t + I_t E_t$.

iv) Employment equals the number of matched workers: $E_t = m(1, v_t)$.

v) Wages are formed through bargaining.

vi) Pooling of after-tax labor income: $I_t = \frac{1}{2}(1-\tau_t) + \frac{1}{2}b$.

vii) Free entry: $k = \frac{m(1,v_t)}{v_t}\frac{1}{2}(1 - \frac{b}{1-\tau_t})$.

viii) Household's asset choice is optimal.
This leads to the consumption Euler equation: $u_c(c_t) = \beta R_{t+1} u_c(c_{t+1})$.

ix) No-Ponzi condition: $\lim_{t\to\infty} q_t B_t = 0$.

I now turn to the characterization of the set of implementable allocations.

[14]Since agreement is immediate in the unique equilibrium of the bargaining game, b only describes the utility derived out of equilibrium. See Hall and Milgrom (2006) for more details and an interpretation of b.

Definition 6. *For any* B_0*, an allocation* $\{c_t, v_t\}_{t=0}^{\infty}$ *is implementable if* $\{\tau_t, w_t, I_t, B_{t+1}, R_{t+1}, E_t\}_{t=0}^{\infty}$ *exists, such that* $\{c_t, g_t = g, v_t, E_t, \tau_t, B_{t+1}, R_{t+1}, I_t, w_t\}_{t=0}^{\infty}$ *is an equilibrium.*

Proposition 3. *For any* B_0*, an allocation* $\{c_t, v_t\}_{t=0}^{\infty}$ *is implementable if and only if*

$$u_c(0)B_0 = \sum_{t\geq 0} \beta^t u_c(t)[c_t - \{\frac{b}{2}\frac{m(1,v_t)/v_t - k}{\frac{1}{2}m(1,v_t)/v_t - k}\}m(1,v_t)]$$
$$c_t + g + kv_t = m(1,v_t).$$

For this model, I can state a result identical to theorem 1, with CS and CU defined as

$$CS(v_t) = u_c(t)[m(1,v_t) - g - kv_t - \{\frac{b}{2}\frac{m(1,v_t)/v_t - k}{\frac{1}{2}m(1,v_t)/v_t - k}\}m(1,v_t)]$$
$$CU(v_t) = u(m(1,v_t) - g - kv_t).$$

Again, I need the concept of time invariance (Definition 4). Here v^{TI} denotes the optimal TI policy, i.e. the constant level of vacancies that is implementable and gives the highest welfare (conditional on being time-invariant). Since further analytical results are not available, I now calibrate the model to check whether condition $(*)$,

$$\frac{CS''}{CU''}(v^{TI}) - \frac{CS'}{CU'}(v^{TI}) < 0,$$

holds. If it does, then the TI-policy v^{TI} is not the Ramsey policy, and a cycle improves welfare.

Government expenditures and government debt B_0 are set equal to zero.[15] The time period is a quarter, thus $\beta = 0.99$. Four parameters remain to be determined: α, χ, k and b.

For α, the parameter of the matching function, Shimer (2005) estimates (consistent with the model) a value of 0.72, using data on the probability of finding a job and on labor market tightness. Other studies (see the survey by Petrongolo and Pissarides (2001)) report lower values between 0.3 and 0.7. Measurement error and

[15] A positive level for B_0 is equivalent to assuming a higher level of government expenditures.

neglecting on-the-job searches seem to render the estimates imprecise. I therefore report results for a range of values for α.

The exact number of vacancies is irrelevant for any result (see Shimer (2005)), as long as it is higher than the number of matches. I thus normalize $v = 1$ (a different value for v only results in different values for χ and k). χ then equals $E \cdot v^{\alpha-1}$ and $k = \frac{1}{2} \cdot (1-b) \cdot E/v$. To determine employment E, Shimer (2005) reports that the monthly unemployment rate averaged 5.67 percent between 1951 and 2003. The efficient solution would of course be no unemployment at all. An economy with an employment rate of 94.33% (1- unemployment rate) is thus quite efficient.[16]

Finally, Hagedorn and Manovskii (2006) find that profits that do exceed 2.2% of revenue are enough to reimburse firms for their vacancy-posting costs.[17] I also tried other values between 1.5% and 3%. The results are insensitive to these variations. In the model, profits equal $\frac{1}{2}(1 - b/(1-\tau))$, which pins down b.

The utility function is assumed to be

$$u(c) \;\; = \;\; c^{1-\sigma}/(1-\sigma).$$

In section 3.3 it was shown that this choice implies no cycles in the LS economy. This result now considerably changes. Table 1 reports, for every value of α, the smallest value for σ, such that a TI policy is not optimal.

Table 1: Parameters for cycles

α	0.3	0.35	0.4	0.45	0.5	0.55	0.6	0.7
σ	1.477	1.602	1.750	1.926	2.141	2.409	2.751	3.829

A low level of inefficiency[18] and wages not formed in competitive markets lead to a

[16]Note that interpreting $1-E$ as nonemployment plus unemployment would render the economy substantially more inefficient and make cycles more likely.

[17]See Hagedorn and Manovskii (2006) for more on this number.

[18]The inefficiency here, stemming from 5.6% of households not working, seems to be small relative to the potential inefficiencies caused by taxation. Prescott (2004) argues that differences in marginal tax rates alone potentially explain why Americans work 50 percent more than do their German, French and Italian counterparts.

situation where cycles are of concern. The size of σ needed in this section is easily within the range that can be found in the macro literature. Gali, Gertler, and Lopez-Salido (2003) and Chari, Kehoe, and McGrattan (2002) all use $\sigma = 5$. McGrattan, Rogerson, and Wright (1997) even estimate values of 5.8 and 6.8 (depending on how they detrend the data).

The reason why cycles are more likely in the search model than in the LS model is that prices/wages are determined differently in equilibrium. In the search model wages change when productivity or the tax rate changes, but do not depend on labor market conditions, such as the number of vacancies or the unemployment rate.[19] In particular, wages remain invariant even when the only (interesting) decision variable in the labor market, the number of vacancies, varies. In the LS model on the other hand, wages are not isolated from the only (interesting) decision variable in the labor market, hours worked. Instead, an increase in hours worked requires an increase in wages. This difference implies that for a two-period tax cycle - where the tax rate is decreased in the first period and increased in the second period - ceteris paribus a larger drop in the tax rate is necessary in the LS model than in the search model to implement the same increase in output. Since this leads to a larger drop in tax revenue in the LS economy in the first period, a larger increase in the tax rate in the second period is necessary to balance the government's budget. This leads to a larger drop in output in the second period, making a tax cycle, in terms of welfare, less attractive in the LS model than in the search model.

The mathematical reason that cycles can improve welfare here lies in the functional form of CS that arises endogenously from the frictions in the model. CS equals $u_c(t)\{c_t - \{\frac{b}{2}\frac{m(1,v_t)/v_t - k}{\frac{1}{2}m(1,v_t)/v_t - k}\}m(1,v_t)\}$ here and equals $u_c(t)c_t - \chi_h(t)h_t$ in the LS economy (using the notation of theorem 2). Whereas $-\chi_h(t)h_t$ is a concave function of c_t and h_t under the assumptions of theorem 2, $-u_c(t)\frac{b}{2}\frac{m(1,v_t)/v_t - k}{\frac{1}{2}m(1,v_t)/v_t - k}m(1,v_t)$ is not

[19] Hall and Milgrom (2006) argue that a realistic threat in a bargaining game between a firm and a worker about a joint surplus is to extend bargaining, and not to terminate it and look for another match. As a consequence, the cost of delay matters for wage determination, not external labor market conditions.

a concave function of c_t and v_t. Therefore the arguments in the proof of theorem 2 do not apply and, as the quantitative exercise above shows, cycles can improve welfare.

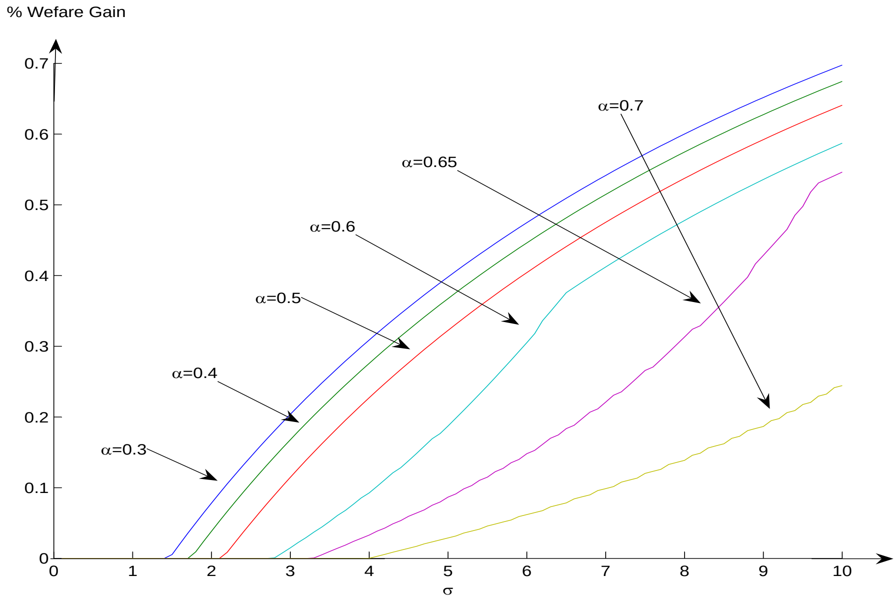
I now use the parameterized model to quantify the welfare effects of implementing a two-period cycle. Theorem 5, which is also applicable in the search and matching model, states that the optimal policy implements only two different levels of vacancies, but not necessarily a two-period cycle. Numerically however I was unable to find a policy that improves upon a two-period cycle. I thus report the welfare gains from implementing such a cycle, bearing in mind that theoretically it is only a lower bound.

The computational strategy is straightforward. I simply compute the welfare for all two-period cycles (for an appropriate grid on the vacancy space). I then choose the cycle which maximizes welfare, and compare it to the welfare from implementing the TI policy v^{TI}. Finally, I express the welfare gain in terms of consumption equivalents. As in Lucas (1987), this measure is defined as the percentage compensation required to make a household indifferent between consumption plans with and without a two-period tax cycle. This number is larger than 0 if a cycle improves welfare, and equals 0 if not.

Figure 2 shows the results for values of σ between 0 and 10 and different values of $\alpha = 0.3, 0.4, 0.5, 0.6, 0.65$, and 0.7. Each graph (for every α) has a kink at the level of σ as computed in Table 1. For all values of σ below this level, there is no welfare gain from a cycle. For higher values of σ, the welfare gains are positive with a maximum level of 0.7%. This is a sizeable number when compared to the cost of business cycles as found by Lucas (1987).

The graphs are ordered by α, the elasticity parameter in the matching function, where a smaller value for α implies larger welfare gains from a cycle. The reason is that with smaller values of α, employment becomes more responsive to changes in vacancies ($E = m(1, v) = \chi v^{1-\alpha}$). For example, in the extreme case of $\alpha = 1$, employment would be constant and policy would be ineffective, whereas employment would be a linear function of vacancies if $\alpha = 0$.

Figure 2: Welfare gains from a two-period cycle in a labor search model



I now show that for cycles to be welfare-improving, it is not essential that frictions arise in the labor market. I therefore consider a model with frictions in the product market in the next section, and show that cycles can improve welfare in this environment as well.

5 Non-Walrasian Product and Credit Markets

In this section I show that a TI policy is not optimal in a parameterized example of the Lagos and Wright (2005) (LW) model. I choose this model since it provides a tractable way to deal with two deviations from the LS economy. First, the product market is not Walrasian. Instead, prices and quantities (of some but not all) goods are determined through bargaining. Second, credit cannot be used to pay for these goods, and in its place a liquid asset, which has to be acquired in advance, is required. Thus in the LW model, both the product market and the credit market have frictions. Since the model does not restrict the interpretation of what this liquid asset is, one can think of it as money (as in LW) or as bonds (as in Lagos (2005)). However, this freedom of interpretation makes a quantitative analysis problematic (see below). I now describe the model and then explain how it fits into the "one-decision" framework of this paper.

Time is discrete and there is a measure one of infinitely-lived agents. Every period is divided into two subperiods. In the first subperiod, agents trade in frictionless markets, followed by a subperiod with trading frictions. I follow the terminology of LW, referring to the frictionless market as the centralized market (CM), and the market with frictions as the decentralized market (DM).

The government buys $g_t = g$ goods in the CM only and has three ways of paying for this: a labor tax τ_t, nominal government bonds B_t, and nominal money creation. The government's flow budget constraint is

$$M_{t+1} + B_{t+1}/R_{t+1} = M_t + B_t + P_t g_t - P_t \tau_t w_t h_t, \tag{14}$$

where M_{t+1} is nominal money outstanding at the end of period t, P_t is the nominal price level in the CM, and R_t is the gross nominal return on bonds. The remaining notation is as in the previous sections.

The household enters the CM in period t with M_t units of money and B_t bonds, and derives utility $W_t(M_t, B_t)$. It decides how much labor h_t to supply, how much to consume c_t and how many bonds B_{t+1} to buy. Production in the CM is, as in LS, a linear function of labor, $F(h_t) = h_t$. In addition to these decisions, which are identical to those taken in the LS economy, every agent can acquire M_{t+1} units of money, which are needed to trade in the subsequent DM. The value of entering the DM with M_{t+1} units of money and B_{t+1} bonds is denoted $V_t(M_{t+1}, B_{t+1})$. The CM problem is

$$W_t(M_t, B_t) = \max_{c_t, h_t, M_{t+1}, B_{t+1}} U(c_t) - Ah_t + V_t(M_{t+1}, B_{t+1}), \tag{15}$$

subject to

$$P_t c_t + B_{t+1}/R_{t+1} + M_{t+1} \leq B_t + M_t + (1 - \tau_t) P_t w_t h_t. \tag{16}$$

Note that LW assume that utility in the CM is linear in labor. This assumption makes the model analytically tractable since (M_{t+1}, B_{t+1}) is independent from (M_t, B_t). In addition, it simplifies the analysis of the DM, since every agent holds the same amount of money.

In the DM, Lagos and Wright (2005) assume that with probability α the household is a buyer, with probability α the household is a seller, and with probability $1-2\alpha$ the household does not trade. A buyer consumes q in the DM and derives utility $u(q)$, while a seller produces q at a cost $c(q)$. u is concave, c is convex and $u(0) = c(0)$ is finite. Then

$$V_t(M_{t+1}, B_{t+1}) = \alpha\{u(q(M_{t+1})) + \beta W_{t+1}(M_{t+1} - d(M_{t+1}), B_{t+1})\} \tag{17}$$

$$+ \alpha\{-c(q(M_{t+1})) + \beta W_{t+1}(M_{t+1} + d(M_{t+1}), B_{t+1})\} \tag{18}$$

$$+ (1-2\alpha)\beta W_{t+1}(M_{t+1}, B_{t+1}), \tag{19}$$

where $q(M_{t+1})$ is the quantity produced by the seller and $d(M_{t+1})$ is the amount paid by the buyer to the seller. This payment has to be in terms of money, $d(M_{t+1}) \leq M_{t+1}$. The terms of trade q and d are determined through Nash bargaining, where the buyer has bargaining power θ. Instead of repeating the derivation in LW, I simply restate the results. Let q^* be defined by $u'(q^*) = c'(q^*)$. Then the solution is

$$q_t(M) = \begin{cases} \hat{q}_t(M) & \text{if } M < M^*_{t+1} \\ q^* & \text{if } M \geq M^*_{t+1}, \end{cases}$$

$$d_t(M) = \begin{cases} M & \text{if } M < M^*_{t+1} \\ M^*_{t+1} & \text{if } M \geq M^*_{t+1}, \end{cases}$$

where $\hat{q}_t(M)$ solves $\beta\chi_t M = z(q_t)$, with

$$z(q) \equiv \frac{\theta c(q)u'(q) + (1-\theta)u(q)c'(q)}{\theta u'(q) + (1-\theta)c'(q)}, \tag{20}$$

and where $M^*_{t+1} = \frac{z(q^*)}{\beta\chi_t}$ and $\chi_t = \frac{A}{P_{t+1}w_{t+1}(1-\tau_{t+1})}$. I can now define an equilibrium.

Definition 7. *Given initial conditions B_0 and M_0, an equilibrium is a sequence $\{c_t, g_t = g, h_t, q_t, d_t, \tau_t, M_{t+1}, B_{t+1}, R_{t+1}, w_t, P_t\}_{t=0}^{\infty}$, such that:*

i) The government budget constraint is satisfied.

ii) The resource constraint in the CM holds: $c_t + g = h_t$.

iii) The household budget constraint (16) is satisfied.

iv) Households choose c_t, h_t, B_{t+1} and M_{t+1} in the CM to maximize utility.

v) Firms in the CM maximize profits.

vi) The terms of trade in the DM are determined through Nash bargaining.

vii) The no-Ponzi condition holds: $\lim_{t\to\infty} q_t B_t = 0$.

In equilibrium it always holds that $M_{t+1} < M^*_{t+1}$. An implementable allocation can be defined as follows:

Definition 8. *For any (B_0, M_0), an allocation $\{c_t, h_t, q_t\}_{t=0}^{\infty}$ is implementable if $\{\tau_t, w_t, d_t, M_{t+1}, B_{t+1}, R_{t+1}, P_t\}_{t=0}^{\infty}$ exists, such that $\{c_t, g_t = g, h_t, q_t, d_t, \tau_t, M_{t+1}, B_{t+1}, R_{t+1}, w_t, P_t\}_{t=0}^{\infty}$ is an equilibrium.*

Using standard techniques (taken from Lucas and Stokey (1983)), Aruoba and Chugh (2006) derive the implementability constraint for this economy and characterize the set of implementable allocations.

Proposition 4. *For any (B_0, M_0), an allocation $\{c_t, h_t, q_t\}_{t=0}^{\infty}$ is implementable if and only if*

$$U_c(c_0)\frac{M_0 + B_0}{P_0} = \sum_{t\geq 0} \beta^t \{U_c(c_t)c_t - Ah_t + \alpha z(q_t)(\frac{u'(q_t)}{z'(q_t)} - 1)\} \tag{21}$$

$$c_t + g = h_t \tag{22}$$

$$\frac{u'(q_t)}{z'(q_t)} \geq 1. \tag{23}$$

As is well known (see for example Chari and Kehoe (1999)), it is necessary to set $M_0 + B_0 = 0$ to make the problem interesting (otherwise the Ramsey planner confiscates the entire initial nominal wealth), and I therefore proceed under this assumption. Since the nominal interest R_{t+1} equals $\alpha(\frac{u'(q_t)}{z'(q_t)} - 1) + 1$, equation (23) is equivalent to a zero bound on nominal interest rates. For the interpretation of $\alpha z(q_t)(\frac{u'(q_t)}{z'(q_t)} - 1)$, note that $z(q_t) = M_{t+1}\beta\chi_t$ and $U_c(c_t) = P_t\chi_{t-1}$ (because utility is quasi-linear in the CM). This implies that $\frac{z(q_t)}{U_c(c_t)} = \frac{M_{t+1}\beta\chi_t}{P_t\chi_{t-1}} = \frac{M_{t+1}}{P_t}\frac{\beta\chi_t}{\chi_{t-1}} = \frac{M_{t+1}}{P_t}\frac{1}{R_{t+1}}$.

Thus $\frac{\alpha z(q_t)(\frac{u'(q_t)}{z'(q_t)}-1)}{U_c(c_t)} = \frac{M_{t+1}}{P_t}\frac{R_{t+1}-1}{R_{t+1}}$, the same expression as in Chari and Kehoe (1999).[20] For a more detailed interpretation, see Aruoba and Chugh (2006).

The Ramsey solution then maximizes

$$\sum_{t\geq 0}\beta^t\{U(c_t) - Ah_t + \alpha(u(q_t) - c(q_t))\} \tag{24}$$

over all implementable allocations. The resource constraint in the CM again implies that c_t is a simple function of h_t, but there is now a second variable, q_t. Nevertheless, the analysis of cycles in the main text can only be applied to the choice of q_t. The same arguments can be used to derive a criterion when the implementation of a time invariant q^{TI} is not optimal (and h is constant and equals the TI solution h^{TI}). The criterion for the choice of q is then identical to that in theorem 1 if

$$CS(q_t) = \alpha z(q_t)(\frac{u'(q_t)}{z'(q_t)} - 1) \quad \text{and} \tag{25}$$

$$CU(q_t) = \alpha(u(q_t) - c(q_t)). \tag{26}$$

I can now evaluate the criterion to assess whether a cycle can improve welfare. Since an analytical evaluation is unavailable, I am obliged to resort to a numerical example. I therefore have to choose functional forms for U, u and c and values for the remaining parameters.

In the CM I set $U(c) = log(c)$, $A = 1/2$, so that cycles in the CM are ruled out by theorem 2. In the DM, the utility function equals $u(q) = \frac{(q+\kappa)^{1-\eta}-\kappa^{1-\eta}}{1-\eta} + 1$, the cost function $c(q) = \exp(q^2/2)$, $\kappa = 1$ and $\alpha = 0.1$. This choice of functional forms would rule out cycles in an LS economy, and cycles would be a rather bad idea since c is very convex. The assumptions needed in LW for the existence of an equilibrium ($u(0) = c(0)$ is finite) make it necessary to add a κ to the utility function and to add 1 to the utility function.

The choice of the functional forms and of α could be guided by two observations in the data, the demand elasticity of M and the level of velocity, i.e. nominal output divided by M. Unfortunately this is somewhat problematic, in particular since these

[20] In this paper $\frac{R_{t+1}-1}{R_{t+1}}$ equals the marginal utility of money divided by U_c.

numbers depend on the interpretation of M in the data. For example, the velocity of bonds is much lower than the velocity of M1. The elasticity of money also depends on the time period. For example, the correlation between $M1/GDP$ and a short-term nominal interest rate is positive for the last 25 years, although it was negative before then. Instead, I use the parameter values as described above, and report the results for the velocity and the demand elasticity.

I vary three parameters, η, the curvature parameter of u, θ, the bargaining power of the buyer and g, government expenditures. For $\theta = 1$, the model with bargaining is equivalent to a model with competitive price-taking in the DM (see Rocheteau and Wright (2005)).

Table 2 reports results and further statistics from my parametrization of Lagos and Wright (2005), such as the size of the CM, the share of government expenditure and the elasticity of money demand. The velocity of money and the elasticity of money demand seem to be reasonable (of course, this statement is subject to the measurement problems discussed above). The size of the government is quite small, between 13% and 17% of CM output.

The time-invariant q is denoted q^{TI}, and q^{Cy} is the constant level of q that makes the representative household indifferent between consumption plans with and without a two-period tax cycle. The percentage gain $\frac{q^{Cy}}{q^{TI}}$ of a cycle in terms of DM consumption is quite large. For example, for $g = 0.4$ and $\eta = 1$ (log preferences), the consumption gain in the DM, $\frac{q^{Cy}}{q^{TI}}$, is close to 50%. The welfare gain can also be expressed in terms of CM consumption. A household is indifferent between 3.97% higher consumption in the CM and a two-period cycle.

Figure 3 provides further information on how the welfare gains from a two-period cycle vary with η. The figure shows $\frac{q^{Cy}}{q^{TI}}$ as a function of η for $\theta = 0.5, 0.75, 1$ and $g = 0.3, 0.4$.

As already observed, a higher value of η unambiguously makes a cycle more likely to be optimal. Changing the value of the buyer's bargaining power θ has two main, potentially opposing, effects on the shape of these curves.

First, θ determines how inefficient q^{TI} is. This affects the criterion $(*)$ mainly

Table 2: Results from the parametrization.

$g = 0.3,\ \eta = 2$						
θ	$\frac{q^{Cy}}{q^{TI}}$	q^{TI}	$\frac{g}{h^{TI}}$	$\frac{DM}{Y=DM+CM}$	$\frac{m^{TI}}{h^{TI}+\alpha m^{TI}}$	$\frac{\partial m}{\partial R}\frac{R}{m}$
0.5	1.225	0.138	0.135	0.024	0.236	-0.236
0.75	1.096	0.199	0.135	0.024	0.236	-0.253
1	1.007	0.272	0.134	0.023	0.234	-0.246

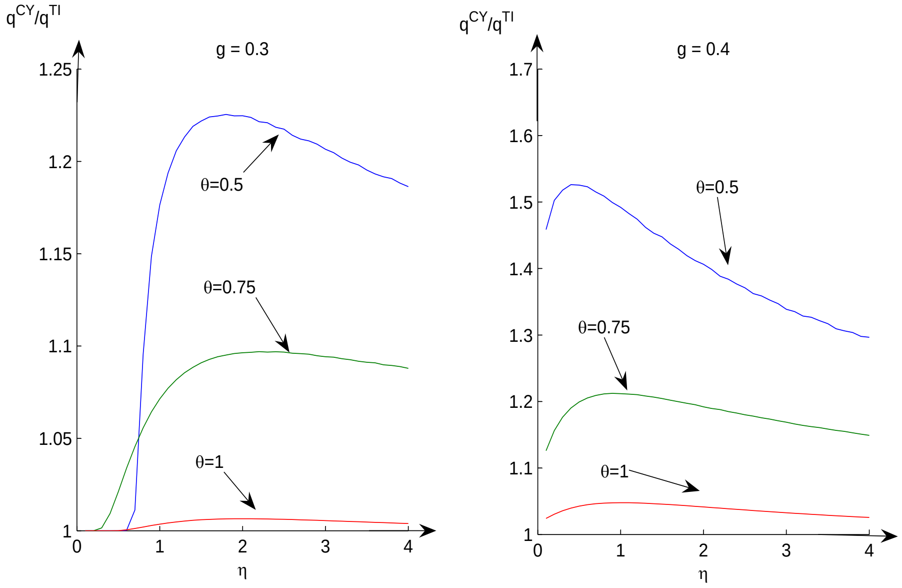
$g = 0.4,\ \eta = 2$						
θ	$\frac{q^{Cy}}{q^{TI}}$	q^{TI}	$\frac{g}{h^{TI}}$	$\frac{DM}{Y=DM+CM}$	$\frac{m^{TI}}{h^{TI}+\alpha m^{TI}}$	$\frac{\partial m}{\partial R}\frac{R}{m}$
0.5	1.406	0.109	0.170	0.022	0.219	-0.125
0.75	1.192	0.165	0.171	0.022	0.221	-0.150
1	1.041	0.235	0.170	0.022	0.221	-0.163

$g = 0.4,\ \eta = 1$						
θ	$\frac{q^{Cy}}{q^{TI}}$	q^{TI}	$\frac{g}{h^{TI}}$	$\frac{DM}{Y=DM+CM}$	$\frac{m^{TI}}{h^{TI}+\alpha m^{TI}}$	$\frac{\partial m}{\partial R}\frac{R}{m}$
0.5	1.492	0.131	0.172	0.023	0.226	-0.199
0.75	1.212	0.206	0.173	0.023	0.231	-0.261
1	1.048	0.292	0.172	0.023	0.231	-0.289

Note: q^{TI}, h^{TI} and c^{TI} are the TI levels of q, h and c. q^{Cy} is the constant level of q that makes the representative household indifferent between consumption plans with and without a two-period tax cycle. $\frac{DM}{Y=DM+CM}$ is the fraction of output Y produced in the DM, $\frac{m^{TI}}{h^{TI}+\alpha m^{TI}}$ is the inverse of money velocity, and $\frac{\partial m}{\partial R}\frac{R}{m}$ is the elasticity of money demand.

through changing CU'. The lower θ is, the smaller the buyer's share of the surplus in the DM. This leads to higher prices and less trade in the DM. Table 2 verifies this claim: q increases in θ, but the amount of money spent does not change substantially (velocity is almost constant). Whether the Ramsey allocation wants to counteract this inefficiency through a lower nominal interest rate (and by how much) depends on the level of g and the standard public finance procedure, i.e. to compare demand elasticities in the DM and in the CM. For $g = 0.3$ and $\theta = 0.5$, the efficient level of q is implemented for small values of η, whereas q^{TI} is always inefficient for $\theta = 0.75, 1$.

Figure 3: Welfare gains from a two-period cycle in Lagos and Wright (2005)



As a result, a cycle does not improve welfare when $\theta = 0.5$ and η is low. However, for higher levels of η, q^{TI} is always lower for a smaller θ (the efficient level of q is independent from θ). Thus cycles become more likely, and the welfare gains become larger if θ is smaller.

The second effect of a change in θ is on the curvature of CS, which means it changes CS'' in the criterion $(*)$. For all parameters, CS'' increases in θ, so that a higher θ makes cycles more likely. Figure 3 then shows the overall effect of θ on both CS'' and CU'. The other two numbers, CU'' and CS', turn out to be of minor importance.

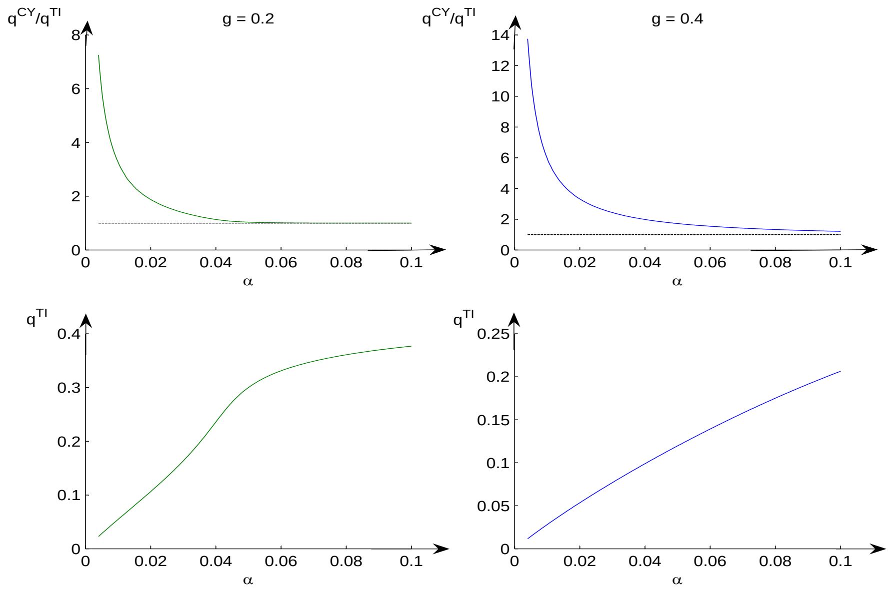
The results so far show that a cycle can improve welfare for all values of η if $g = 0.4$ and even if $\theta = 1$ (competitive pricing). The reason that functional forms, which would imply the optimality of a TI policy in the LS economy, lead to welfare-improving cycles, even for competitive pricing, is because of credit market frictions. The credit market frictions which arise endogenously in LW lead to a functional form for CS, which differs from approaches where money is an argument of the utility function (Chari and Kehoe (1999)). This has two consequences. First, the Friedman rule is not necessarily optimal, which implies that q^{TI} is not efficient even if $\theta = 1$.

This is a necessary condition for a cycle, since efficiency means $CU' = 0$ and the criterion $(*)$ cannot be fulfilled. The second implication is that the credit market friction adds an additional nonlinearity to the function that maps quantities into tax revenue (CS). In the labor search model and also in LW if $\theta < 1$, bargaining leads to a nonlinear deviation from the Walrasian mapping of quantities into prices, and thus mechanically also changes the mapping from quantities into tax revenue. On top of that, only the medium of exchange is taxed in LW and not trading, which adds some nonlinearity to the mapping from quantities to tax revenues even if pricing is competitive (CS'' becomes positive).

To analyze how the size of the friction affects the likelihood of an optimal tax cycle, I consider now the experiment of decreasing α, i.e. the probability that money is needed in the DM. A decrease in α can be interpreted as an increase in the probability ρ that credit can be used in a DM meeting. In the benchmark, $\rho = 0$ and an agent is a buyer, needing money with probability $\alpha = 0.1$. A value of $\alpha = 0.05$ then means that with probability 0.1 an agent is a buyer, but needs money only in 50% of these meetings and can pay with credit otherwise. If credit frictions are fully removed ($\alpha = 0$), money is not needed and credit can be used in all transactions. In this limiting case the economy is identical to an LS economy, and a TI policy is optimal in both the CM and the DM.

Figure 4 shows the result for log preferences in the DM, a bargaining power of $\theta = 0.75$ and two different levels of government expenditures $g = 0.2, 0.4$. For both levels of g Figure 4 shows the percentage welfare gain $\frac{q^{Cy}}{q^{TI}}$ and q^{TI} as a function of α. Somewhat surprisingly, reducing frictions (reducing α) increases $\frac{q^{Cy}}{q^{TI}}$. Furthermore, for $g = 0.2$, condition $(*)$ does not hold for high values of α, although it does hold for low values of α. The plot of q^{TI} explains both results. As α decreases, inefficiencies in the DM have smaller welfare effects and thus the Ramsey solution implements a lower q^{TI}. Given that a cycle improves welfare, a lower q^{TI} (mechanically) increases $\frac{q^{Cy}}{q^{TI}}$, since the denominator becomes a small number. A cycle is also more likely to be welfare-improving as explained above. The less efficient the economy becomes (the lower q^{TI}), the higher CU' is, making it then more likely that condition $(*)$ holds.

Figure 4: Welfare gains from a two-period cycle in Lagos and Wright (2005) as a function of α (credit market friction)



But inefficiency alone does not imply cycles as I showed in the LS economy (theorem 2 does not depend on the level of g). The second requirement, convexity of CS, is however unaffected by changes in α. The same convex functional form describes the government's tax revenue from the DM.

6 Conclusion

I consider whether tax smoothing or cycles represent the solution to a Ramsey taxation problem both in a Lucas-Stokey economy and in two models with frictions. There is no extrinsic uncertainty that would result in cycles.

In an LS economy, a large class of preferences exists where cycles are optimal. I show that sufficiently strong DRRA implies that cycles are optimal, whereas IRRA implies that they are not.

I then consider two simple models with frictions: a labor search and matching economy, following Pissarides (2000); and the Lagos and Wright (2005) economy, a model with frictions in the goods and credit market. A cycle is now optimal for a class of preferences that was considered to imply the opposite in the Lucas-Stokey world. Thus the results established for the Lucas-Stokey economy, which guaranteed optimal tax smoothing for a certain class of preferences, do not prove to carry over to all models of interest.

My results suggest the possibility that cycles could improve welfare. This is especially true in models with frictions, which invalidate the first welfare theorem and lead to noncompetitive pricing.[21] For specific applications, the theorems in this paper are helpful along two dimensions. First, they can imply that cycles are never welfare-improving, so that one can proceed with a first-order approach. Second, they provide a condition to check whether a cycle can improve welfare. Although this condition has to be checked in every application, it provides some guidance to the features that potentially lead to cycles: inefficient allocations, and non linear pricing that leads to a convex tax revenue function.

Proofs

Proof of Theorem 1

In principle, four cases have to be considered. Labor supply can be too high ($CU' < 0$) or too low ($CU' > 0$), and tax revenue can be increasing ($CS' > 0$) or decreasing ($CS' < 0$) in h. Two cases can be ruled out immediately if we allow for transfers from the government to the household (this means the implementability constraint becomes an inequality constraint; in any case I use this "trick" later.). If $CU' > 0$ and $CS' > 0$, an increase in labor supply is welfare-improving and raises tax revenues. Conversely, if $CU' < 0$ and $CS' < 0$, a decrease in labor supply is welfare-improving and raises tax revenues. Two cases remain. In Case I, the labor supply is too low ($CU' > 0$) and cutting taxes decreases tax revenue ($CS' < 0$). In Case II, labor supply is too high ($CU' < 0$) and the economy is on the downside of the Laffer curve, so that lowering taxes increases tax revenue ($CS' > 0$). In Case I ($*$) is

[21] Adding capital to these models (and allowing for capital income taxation) would change the quantitative results but not the main conclusions of the paper.

equivalent to

$$\frac{CS''}{CS'}(h^{TI}) - \frac{CU''}{CU'}(h^{TI}) < 0 \qquad (*^I)$$

and in Case II $(*)$ is equivalent to

$$\frac{CS''}{CS'}(h^{TI}) - \frac{CU''}{CU'}(h^{TI}) > 0 \qquad (*^{II})$$

I use a perturbation argument to show that a TI policy is not optimal. Consider two consecutive periods (t and $t+1$) where the same amount of labor h^{TI} is implemented. In the first period, labor supply is increased by $\Delta \geq 0$. In the second period, labor supply is decreased by $\delta/\beta \geq 0$. Let $S(\Delta, \delta)$ be the present value surplus and $V(\Delta, \delta)$ the present value utility from these two periods (evaluated at h^{TI}). Thus

$$\begin{aligned} S(\Delta, \delta) &= CS(h^{TI} + \Delta) + \beta \cdot CS(h^{TI} - \delta/\beta) \\ V(\Delta, \delta) &= CU(h^{TI} + \Delta) + \beta \cdot CU(h^{TI} - \delta/\beta) \end{aligned}$$

For small Δ, define $\delta^S(\Delta)$ and $\delta^V(\Delta)$ such that

$$\begin{aligned} S(\Delta, \delta^S(\Delta)) &= 0 \\ V(\Delta, \delta^V(\Delta)) &= 0 \end{aligned}$$

If multiple solutions exist, the smallest δ is chosen. Define $\delta^*(\Delta) = (\delta^S(\Delta) + \delta^V(\Delta))/2$.

If condition $(*)$ holds, it will be shown that for small Δ, a joint increase by Δ in the first period and a decrease by $\delta(\Delta)/\beta$ in the second period increase both S and V. An increase in S means that this policy change is implementable by proposition 14, while an increase in V means that this policy change is welfare-improving.

Thus let $\tilde{S}$ and $\tilde{B}$ be defined as follows:

$$\begin{aligned} \tilde{S}(\Delta) &= S(\Delta, \delta^*(\Delta)) \\ \tilde{V}(\Delta) &= V(\Delta, \delta^*(\Delta)) \end{aligned}$$

Obviously $\delta^*(0) = (\delta^S(0) + \delta^V(0))/2 = 0$. Implicit differentiation results in:

$$\begin{aligned}
\frac{\partial\delta^*(\cdot)}{\partial\Delta}(0) = \frac{\partial\delta^S(\cdot)}{\partial\Delta}(0) = \frac{\partial\delta^V(\cdot)}{\partial\Delta}(0) &= 1 \\
\frac{\partial^2\delta^S(\cdot)}{\partial\Delta\partial\Delta}(0) &= \frac{(1+\beta)\cdot CS''(h^{TI})}{\beta\cdot CS'(h^{TI})} \\
\frac{\partial^2\delta^V(\cdot)}{\partial\Delta\partial\Delta}(0) &= \frac{(1+\beta)\cdot CU''(h^{TI})}{\beta\cdot CU'(h^{TI})}.
\end{aligned}$$

Consider first derivatives with respect to Δ:

$$\begin{aligned}
\frac{\partial\tilde{S}}{\partial\Delta} &= CS'(h^{TI}+\Delta) - CS'(h^{TI} - \delta^*(\Delta)/\beta)\cdot\delta^{*'}(\Delta) \\
&\overset{\Delta=0}{=} 0 \\
\frac{\partial\tilde{V}}{\partial\Delta} &= CU'(h^{TI}+\Delta) - CU'(h^{TI} - \delta^*(\Delta)/\beta)\cdot\delta^{*'}(\Delta) \\
&\overset{\Delta=0}{=} 0.
\end{aligned}$$

This shows that a linear approximation does not produce an affirmative answer. Therefore the sign of the second derivatives is decisive:

$$\begin{aligned}
\frac{\partial^2\tilde{S}}{\partial\Delta\partial\Delta} &= CS''(h^{TI}+\Delta) + CS''(h^{TI} - \delta^*(\Delta)/\beta)/\beta\cdot(\delta^{*'}(\Delta))^2 - CS'(h^{TI} - \delta^*(\Delta)/\beta)\cdot\delta^{*''}(\Delta) \\
&\overset{\Delta=0}{=} CS''(h^{TI})(1+\beta)/\beta - CS'(h^{TI})\cdot\delta^{*''}(0) \\
&= (1+\beta)/\beta\cdot\{CS''(h^{TI}) - CS'(h^{TI})/2\cdot[\frac{CS''(h^{TI})}{CS'(h^{TI})} + \frac{CU''(h^{TI})}{CU'(h^{TI})}]\} \\
&> (1+\beta)/\beta\cdot\{CS''(h^{TI}) - CS'(h^{TI})/2\cdot[2\cdot\frac{CS''(h^{TI})}{CS'(h^{TI})}]\} \\
&= 0,
\end{aligned}$$

where the inequality sign follows in case I from $(*^I)$ and $CS'(h^{TI}) < 0$, and in case II from $(*^{II})$ and $CS'(h^{TI}) > 0$. The same calculations for V result in:

$$\begin{aligned}
\frac{\partial^2\tilde{V}}{\partial\Delta\partial\Delta} &\overset{\Delta=0}{=} CU''(h^{TI})(1+\beta)/\beta - CU'(h^{TI})\cdot\delta^{*''}(0) \\
&= (1+\beta)/\beta\cdot\{CU''(h^{TI}) - CU'(h^{TI})/2\cdot[\frac{CS''(h^{TI})}{CS'(h^{TI})} + \frac{CU''(h^{TI})}{CU'(h^{TI})}]\} \\
&> (1+\beta)/\beta\cdot\{CU''(h^{TI}) - CU'(h^{TI})/2\cdot[2\cdot\frac{CU''(h^{TI})}{CU'(h^{TI})}]\} \\
&= 0,
\end{aligned}$$

where the inequality sign again follows in case I from $(*^I)$ and $CU'(h^{TI}) > 0$, and in case II from $(*^{II})$ and $CU'(h^{TI}) < 0$.

Proof of Theorem 2

The Ramsey problem is to choose c_t and h_t to maximize utility subject to implementability and resource constraints. In a first step it is shown that both these constraints can be written as inequalities, such that the (relaxed) Ramsey problem reads as follows:

$$\begin{aligned} Max_{c_t,h_t} \quad & \sum_{t=0}^{\infty} \beta^t (w(c_t) - \chi(h_t)) \\ s.t. \quad & \sum_{t\geq 0} \beta^t (w_c(c_t) \cdot c_t - \chi_h(h_t) \cdot h_t) \geq 0 \\ & c_t + g \leq h_t. \end{aligned}$$

Suppose first that in the (relaxed) maximization problem, the resource constraint is fulfilled with strict inequality for some t. Then decreasing h_t improves welfare and still satisfies the relaxed implementability constraint (the implementability constraint is decreasing in h_t).

Thus the resource constraint is fulfilled with equality for all t, and $c_t = h_t - g$.

Suppose now that the implementability constraint is fulfilled with strict inequality. Then for some t:

$$(w_c(c_t) \cdot c_t - \chi_h(h_t) \cdot h_t) > 0.$$

Since $c_t + g = h_t$,:

$$\begin{aligned} 0 &< (w_c(c_t) \cdot c_t - \chi_h(h_t) \cdot h_t) \\ &= w_c(c_t) \cdot (h_t \cdot (1 - \chi_h(h_t)/w_c(c_t)) - g) \\ &\Rightarrow (1 - \chi_h(h_t)/w_c(c_t)) > 0 \\ &\Leftrightarrow w_c(c_t) > \chi_h(h_t). \end{aligned}$$

Thus increasing h_t and c_t by the same (small) amount improves welfare. The relaxed implementability constraint is still fulfilled, since it was assumed to hold with strict inequality.

These arguments imply that the solution to the relaxed problem fulfills both constraints with equality. Therefore the solution to the relaxed problem and the solution to the Ramsey problem coincide.

A substitution of variables $v_t := w(c_t)$ results in an equivalent maximization problem (P^*) which reads as follows:

$$\begin{array}{ll} Max_{v_t,h_t} & \sum_{t=0}^{\infty} \beta^t (v_t - \chi(h_t)) \\ s.t. & \sum_{t \geq 0} \beta^t (\varphi(\psi(v_t)) - \chi_h(h_t) \cdot h_t) \geq 0 \\ & \psi(v_t) + g \leq h_t \end{array}$$

for all feasible v_t (those where a c_t exists such that $v_t = u(c_t)$, $\psi(v) = w^{-1}(v)$ and $\varphi(c) = w_c(c) \cdot c$.

It will be shown that this problem is convex for IRRA utility functions. This amounts to showing that both constraints are convex.

This is true for the resource constraint, since the inverse of any increasing concave utility function is convex.

For the implementability constraint, $\kappa(v) := \varphi(\psi(v_t))$ has to be a concave function (concave since the inequality sign is $\geq$ and not $\leq$.). $\chi_h(h_t) \cdot h_t$ is a convex function by assumption, so that

$$\begin{aligned} \kappa''(v) &= \varphi''(\psi')^2 + \varphi'\psi'' \\ &= \frac{2w'' + w''' \cdot c}{(w')^2} - \frac{(w' + w''c)w''}{(w')^3} \\ &= \frac{w''}{(w')^2}(1 + \frac{w'''c}{w''} - \frac{w''c}{w'}) \leq 0 \end{aligned}$$

since w is increasing and concave and IRRA is equivalent to

$$1 + \frac{w'''c}{w''} - \frac{w''c}{w'} \geq 0.$$

Since problem (P^*) is strictly concave, first-order conditions for every t have a unique solution (v_t^*, h_t^*). Since first-order conditions are necessary for an optimum, (v_t^*, h_t^*) is the optimal choice at t. Furthermore, since the maximization problem is identical for all t, the same (v^*, h^*) is implemented in every period t. As problem (P^*) is equivalent to the Ramsey problem, $(c^* := w^{-1}(v^*), h^*)$ is the optimal choice in every period.

Proof of Theorem 3

This follows from theorem 4.

Proof of Theorem 4

I check condition (*) of theorem 2. The functional form of the utility function implies that

$$\begin{aligned} CU' &= w' - \chi' \\ CU'' &= w'' - \chi'' \\ CS' &= w' + w''c - \chi' - \chi''h \\ CS'' &= 2w'' + w'''c - 2\chi'' - \chi'''h \\ \frac{\chi'}{w'} &= 1 - \tau^{TI} \end{aligned}$$

and therefore

$$\begin{aligned} \frac{CS'}{CU'} &= 1 + \frac{w''c - \chi''h}{w' - \chi'} = 1 + \frac{w''c}{w' - \chi'} - \frac{\chi''h}{w' - \chi'} \\ &= 1 + \frac{w''c}{w'\tau^{TI}} - \frac{\chi''h}{\chi'\frac{\tau^{TI}}{1-\tau^{TI}}} = 1 + \eta_c\frac{1}{\tau^{TI}} - \eta_h\frac{1-\tau^{TI}}{\tau^{TI}} \\ \frac{CS''}{CU''} &= 2 + \frac{w'''c - \chi'''h}{w'' - \chi''} = 2 + \frac{w'''c}{w'' - \chi''} - \frac{\chi'''h}{w'' - \chi''} \\ &= 2 + \frac{w'''c}{w''}\frac{1}{1 - \frac{\chi''}{w''}} - \frac{\chi'''h}{\chi''}\frac{1}{\frac{w''}{\chi''} - 1}. \end{aligned}$$

Since

$$\frac{\chi''}{w''} = \eta_h\frac{\chi'}{w''h} = \eta_h\frac{w'(1-\tau^{TI})^2}{w''c} = \frac{\eta_h(1-\tau^{TI})^2}{\eta_c}$$

it follows that the optimality criterion

$$\begin{aligned} \frac{CS''}{CU''} - \frac{CS'}{CU'} &= 1 + \kappa_c\frac{1}{1 - \frac{\eta_h(1-\tau^{TI})^2}{\eta_c}} - \kappa_h\frac{1}{\frac{\eta_c}{(1-\tau^{TI})^2\eta_h} - 1} - \eta_c\frac{1}{\tau^{TI}} + \eta_h\frac{1-\tau^{TI}}{\tau^{TI}} \\ &= 1 + \frac{\kappa_c\eta_c - \kappa_h\eta_h(1-\tau^{TI})^2}{\eta_c - \eta_h(1-\tau^{TI})^2} + \frac{\eta_h(1-\tau^{TI}) - \eta_c}{\tau^{TI}} < 0. \end{aligned}$$

Proof of Proposition 2

The assumption that $\kappa_h \geq -2$ implies that

$$\begin{aligned} \frac{CS'}{CU'} &= 1 + \eta_c\frac{1}{\tau^{TI}} - \eta_h\frac{1-\tau^{TI}}{\tau^{TI}} \leq 1 + \eta_c\frac{1}{\tau^{TI}} \\ \frac{CS''}{CU''} &= \frac{2w'' + w'''c}{w'' - \chi''} - \frac{2\chi'' + \chi'''h}{w'' - \chi''} = \frac{w''}{w'' - \chi''}(2 + \kappa_c) - \frac{\chi''}{w'' - \chi''}(2 + \kappa_h) \geq \frac{w''}{w'' - \chi''}(2 + \kappa_c). \end{aligned}$$

If $2+\kappa_c \geq 0$, then $\frac{CS''}{CU''} \geq 0$. Since $\frac{CS'}{CU'} \leq 0$ at the optimal TI policy (see the arguments in the proof of theorem 1), condition $(***)$ is not satisfied.

If $2+\kappa_c \leq 0$, then $\eta_c < 0$ and

$$\frac{CS''}{CU''} - \frac{CS'}{CU'} \;\geq\; \frac{w''}{w''-\chi''}(2+\kappa_c) - (1+\frac{\eta_c}{\tau^{TI}}) \geq (2+\kappa_c) - (1+\frac{\eta_c}{\tau^{TI}}).$$

Both claims follow since $(2+\kappa_c)-(1+\frac{\eta_c}{\tau^{TI}}) \geq (2+\kappa_c)-(1+\eta_c)$, which is positive for IRRA preferences.

Proof of Theorem 5

For a fixed set of labor supply levels $\{h_1,\ldots h_n\}$, consider the following convexified problem:

$$\max_{\pi=(\pi_1,\ldots,\pi_n)} \sum_{i=1}^{n} \pi_i v(h_i) \tag{27}$$

subject to

$$\sum_{i=1}^{n} \pi_i \gamma(h_i) = 0 \tag{28}$$

$$\sum_{i=1}^{n} \pi_i = 1 \tag{29}$$

$$0 \leq \pi_i \leq 1, \tag{30}$$

where $v(h_t) = u(h_t - g, h_t)$ and $\gamma(h_t) = u_c(h_t-g,h_t)\cdot(h_t-g) + u_h(h_t-g,h_t)\cdot h_t$. A π is called *admissible* if it fulfills (28)-(30). It is a standard result in linear programming (see e.g. Dantzig (1963)) that a corner (or vertex or extreme point) solution is optimal, where a corner is defined as follows:

Definition 9. *An admissible $\pi = \{\pi_1,\ldots,\pi_n\}$ is a corner (or vertex or extreme point) if for all admissible $\tilde{\pi} = \{\tilde{\pi}_1,\ldots,\tilde{\pi}_n\}$ and $\hat{\pi} = \{\hat{\pi}_1,\ldots,\hat{\pi}_n\}$ and $\lambda \in (0,1)$:*

$$\pi = \lambda\tilde{\pi} + (1-\lambda)\hat{\pi} \;\;\Rightarrow\;\; \pi = \tilde{\pi} = \hat{\pi}. \tag{31}$$

I show next that for a corner π, at most two π_i are not zero. Suppose there is a corner with a support larger than two. Thus there exist three different i (i_1, i_2, i_3) such that $\pi_{i_1}, \pi_{i_2}, \pi_{i_3} \neq 0$. I now define $\tilde{\pi} \neq \pi$ and $\hat{\pi} \neq \pi$ such that $\frac{1}{2}\tilde{\pi} + \frac{1}{2}\hat{\pi} = \pi$, i.e.

π is not a corner. Note that without loss of generality all $\gamma_j := \gamma(h_{i_j})$ $(j \in \{1,2,3\})$ take different values, since for the same γ value, only the h that gives the highest utility is chosen. Define

$$\tilde{\pi}_{i_1} := \pi_1 - \epsilon \frac{\gamma_2 - \gamma_3}{\gamma_1 - \gamma_3} \tag{32}$$
$$\tilde{\pi}_{i_2} := \pi_2 + \epsilon \tag{33}$$
$$\tilde{\pi}_{i_3} := \pi_3 - \epsilon \frac{\gamma_1 - \gamma_2}{\gamma_1 - \gamma_3} \tag{34}$$
$$\hat{\pi}_{i_1} := \pi_1 + \epsilon \frac{\gamma_2 - \gamma_3}{\gamma_1 - \gamma_3} \tag{35}$$
$$\hat{\pi}_{i_2} := \pi_2 - \epsilon \tag{36}$$
$$\hat{\pi}_{i_3} := \pi_3 + \epsilon \frac{\gamma_1 - \gamma_2}{\gamma_1 - \gamma_3} \tag{37}$$

for some sufficiently small $\epsilon > 0$. For all other indices, $\tilde{\pi}, \hat{\pi}$ and π coincide.
Since $\tilde{\pi}$ and $\hat{\pi}$ are admissible and fulfill $\frac{1}{2}\tilde{\pi} + \frac{1}{2}\hat{\pi} = \pi$, π is not a corner.
Thus the solution of the convexified problem has the desired property of a support not larger than two.
I now show that this is also the case for the Ramsey problem. Suppose the Ramsey problem implements labor supply levels $h^1, \ldots h^n$, where h^i is implemented at dates $I_i := \{t \mid h_t = h^i\}$. Define π as follows:

$$\pi_i := (1-\beta) \sum_{t \in I_i} \beta^t \tag{38}$$

Since the Ramsey solution fulfills the implementability constraint, π is admissible. The above arguments show that an admissible $\tilde{\pi}$ exists with a support not larger than two, which produces the same welfare level. This implies that

$$(1-\beta) \sum_{t=0}^{\infty} \beta^t v(h_t) = \sum_{i=1}^{n} \pi_i v(h^i) \leq \tilde{\pi}_1 v(h^{i_1}) + \tilde{\pi}_2 v(h^{i_2}) \tag{39}$$

for some $1 \leq i_1, i_2 \leq n$.
I now show that $J_1, J_2 \subset \mathbb{N}_0$ exist, such that $J_1 \cup J_2 = \mathbb{N}_0$ and

$$(1-\beta) \sum_{s \in J_1} \beta^s = \tilde{\pi}_1 \quad \text{and} \tag{40}$$
$$(1-\beta) \sum_{s \in J_2} \beta^s = \tilde{\pi}_2 \tag{41}$$

The following algorithm achieves this. Define a sequence of index sets $J_1(t), J_2(t)$ as follows: $J_1(-1) = J_2(-1) = \emptyset$.

$$J_1(t) = J_1(t-1) \cup \{t\}, J_2(t) = J_2(t-1), \quad \text{if } \tilde{\pi}_1 - \hat{\pi}_1(t-1) \geq \tilde{\pi}_2 - \hat{\pi}_1(t-1)$$
$$J_1(t) = J_1(t-1), J_2(t) = J_2(t-1) \cup \{t\}, \quad \text{otherwise,}$$

where $\hat{\pi}_i(t) := (1-\beta)\sum_{s\in J_i(t)} \beta^s$ and $\hat{\pi}_i(-1) = 0$.
$J_1 := \lim_{t\to\infty} J_1(t)$ and $J_2 := \lim_{t\to\infty} J_2(t)$.
This algorithm works if $\tilde{\pi}_i \geq \hat{\pi}_i(t)$ holds for all $t \geq 0$. This is true for t if it is true for $t-1$ and if

$$(1-\beta)\beta^t \quad \leq \quad \max\{\tilde{\pi}_1 - \hat{\pi}_1(t-1), \tilde{\pi}_2 - \hat{\pi}_2(t-1)\}. \tag{42}$$

Since $\tilde{\pi}_1 + \tilde{\pi}_2 = 1$, $\hat{\pi}_1(t-1) + \hat{\pi}_2(t-1) = (1-\beta)\sum_{s=0}^{t-1} \beta^s = 1 - \beta^t$ and $\beta > 1/2$

$$\tilde{\pi}_1 + \tilde{\pi}_2 - \hat{\pi}_1(t-1) - \hat{\pi}_2(t-1) = \beta^t > 2(1-\beta)\beta^t. \tag{43}$$

The algorithm converges because (43) implies (42) and $\hat{\pi}_i(t) \to \tilde{\pi}_i$. Thus the same welfare level can be reached with two different tax levels:

$$(1-\beta)\sum_{t=0}^{\infty} \beta^t v(h_t) = \sum_{i=1}^{n} \pi_i v(h^i) \tag{44}$$
$$\leq \quad \tilde{\pi}_1 v(h^{i_1}) + \tilde{\pi}_2 v(h^{i_2}) = (1-\beta)(\sum_{s\in J_1} \beta^s v(h^{i_1}) + \sum_{s\in J_2} \beta^s v(h^{i_2}))$$

An infinite support can be ruled out as follows. Consider the problem truncated at time n. The support of these problems is not larger than two. Since the welfare of the solution for $n = \infty$ is the limit for $n \to \infty$, the optimal solution can always be approximated by a solution with a finite support (of size two).

Proof of Proposition 3

The household's present value budget constraint reads as follows:

$$B_0 \quad = \quad \sum_{t\geq 0} q_t \cdot \{c_t - I_t \cdot E_t\} \tag{45}$$

From the free-entry condition $k = \frac{m(1,v_t)}{v_t}\frac{1}{2}(1 - \frac{b}{1-\tau_t})$, it follows that

$$(1-\tau_t) \quad = \quad \frac{\frac{b}{2}m(1,v_t)/v_t}{\frac{1}{2}m(1,v_t)/v_t - k}. \tag{46}$$

Income then equals

$$I_t = \frac{1}{2}(1-\tau_t) + \frac{1}{2}b \tag{47}$$

$$= \frac{1}{2}\frac{\frac{b}{2}m(1,v_t)/v_t}{\frac{1}{2}m(1,v_t)/v_t - k} + \frac{1}{2}b \tag{48}$$

$$= \frac{b}{2}\{1 + \frac{\frac{1}{2}m(1,v_t)/v_t}{\frac{1}{2}m(1,v_t)/v_t - k}\} \tag{49}$$

$$= \frac{b}{2}\{\frac{m(1,v_t)/v_t - k}{\frac{1}{2}m(1,v_t)/v_t - k}\}. \tag{50}$$

The proposition follows since the consumption Euler equation implies that $q_t = \beta^t \cdot \frac{u_c(t)}{u_c(0)}$ and $E_t = m(1, v_t)$.

References

Aiyagari, R. S. (1994): "Optimal Capital Income Taxation with Incomplete Markets, Borrowing Constraints, and Constant Discounting," Working Paper No 508, Federal Reserve Bank of Minneapolis.

Aruoba, B., and S. Chugh (2006): "Optimal Fiscal and Monetary Policy When Money Is Essential," mimeo.

Barro, R. J. (1979): "On the Determnation of the Public Debt," *Journal of Political Economy*, 87, 940–971.

Binmore, K., A. Rubinstein, and A. Wolinsky (1986): "The Nash Bargaining Solution in Economic Modeling," *Rand Journal of Economics*, 17(2), 176–188.

Boldrin, M., and L. Montrucchio (1995): "Acyclicity and Dynamic Stability: Generalizations and Applications," *Journal of Economic Theory*, 65, 303–326.

Chamley, C. (1986): "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives," *Econometrica*, 54(3), 607–622.

Chari, V. V., and P. Kehoe (1999): "Optimal Fiscal and Monetary Policy," in *Handbook of Macroeconomics*, ed. by J. Taylor, and M. Woodford, vol. 1. Amsterdam: North-Holland.

Chari, V. V., P. Kehoe, and E. McGrattan (2002): "Can Sticky Price Models Generate Volatile and Persistent Real Exchange Rates?," *Review of Economic Studies*, 69, 533–563.

Coles, M. (2006): "Optimal Unemployment Insurance with Hidden Search Effort and Endogenous Savings," mimeo, University of Essex.

Gali, J., M. Gertler, and J. Lopez-Salido (2003): "Markups, Gaps, and the Welfare Costs of Business Fluctuations," NBER Working Paper W8850.

Gollier, C. (2001): *The Economics of Risk and Time*. MIT press, Cambridge, MA.

Hagedorn, M., and I. Manovskii (2006): "The Cyclical Behavior of Equilibrium Unemployment and Vacancies Revisited," mimeo, University of Pennsylvania.

Hall, R. E. (1988): "Intertemporal Substitution in Consumption," *Journal of Political Economy*, 96(2), 339–357.

Hall, R. E., and P. R. Milgrom (2006): "The Limited Influence of Unemployment on the Wage Bargain," mimeo, Stanford University.

Hansen, G. (1985): "Indivisible Labor and the Business Cycle," *Journal of Monetary Economics*, 16, 309–327.

Hassler, J., P. Krusell, K. Storesletten, and F. Zilibotti (2004): "On the Optimal Timing of Capital Taxes," CEPR Discussion Paper 4731.

Judd, K. (1985): "Redistributive Taxation in a Simple Perfect Foresight Model," *Journal of Public Economics*, 28(1), 59–83.

Kocherlakota, N. R. (2004): "Figuring out the Impact of Hidden Savings on Optimal Unemployment Insurance," *Review of Economic Dynamics*, 7(3), 541–554.

——— (2006): "Advances in Dynamic Optimal Taxation," in *Advances in Economics and Econometrics: Theory and Applicatons, Ninth World Congress*, vol. I, pp. 269–299.

Lagos, R. (2005): "Asset Prices and Liquidity in an Exchange Economy," mimeo, New York University.

Lagos, R., and R. Wright (2005): "A Unified Framework for Monetary Theory and Policy Analysis," *Journal of Political Economy*, 113(3), 463–484.

Lucas, R. E. J. (1987): *Models of Business Cycles*. Basil Blackwell, New York.

Lucas, R. E. J., and N. L. Stokey (1983): "Optimal Fiscal and Monetary Policy in an Economy without Capital," *Journal of Monetary Economics*, 12(1), 55–93.

McGrattan, E., R. Rogerson, and R. Wright (1997): "An Equilibrium Model of the Business Cycle with Household Production and Fiscal Policy," *International Economic Review*, 38(2), 267–290.

McKenzie, L. (1986): "Optimal Economic Growth, Turnpike Theorems and Comparative Dynamics," in *Handbook of Mathematical Economics*, ed. by K. Arrow, and M. Intriligator, vol. III. Amsterdam: North-Holland.

Petrongolo, B., and C. Pissarides (2001): "Looking into the Black Box: A Survey of the Matching Function," *Journal of Economic Literature*, 39, 390–431.

Pissarides, C. (1985): "Short-run Equilibrium Dynamics of Unemployment, Vacancies and Real Wages," *American Economic Review*, 75, 676–690.

Pissarides, C. (2000): *Equilibrium Unemployment Theory*. MIT Press, Cambridge, MA.

Prescott, E. C. (2004): "Why Do Americans Work So Much More Than Europeans," *Federal Reserve Bank of Minneapolis Quarterly Review*, 28(1), 2–13.

Rocheteau, G., and R. Wright (2005): "Money in Search Equilibrium, in Competitive Equilibrium, and in Competitive Search Equilibrium," *Econometrica*, 74, 193–212.

Rogerson, R. (1988): "Indivisible Labor, Lotteries and Equilibrium," *Journal of Monetary Economics*, 21, 3–16.

Shimer, R. (2005): "The Cyclical Behavior of Equilibrium Unemployment and Vacancies," *American Economic Review*, 95(1), 25–49.

Shimer, R., and L. Smith (2001): "Nonstationary Search," mimeo.

Vissing-Jorgenson, A. (2002): "Limited Asset Market Participation and the Elasticity of Intertemporal Substitution," *Journal of Political Economy*, 110, 825–853.

Woodford, M. (2003): *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton, NJ: University Press.