Supplement to "Optimal fiscal policy with heterogeneous agents"

(*Quantitative Economics*, Vol. 5, No. 3, November 2014, 675–704)

Marco Bassetto
University College London, Federal Reserve Bank of Chicago, and IFS

Appendix D: Proof of Theorem 1

Since the proof of Theorem 1 is rather cumbersome, it is useful to break it into several lemmas.

We first rewrite the objective function in terms of the integrals I_1, I_2, $\hat{I}_2$, I_3, and I_4, defined as

$$I_1 \equiv \int e(v)^{1-\gamma}\, dm(v), \tag{58}$$

$$I_2 \equiv \int \bigl(Mb^1(v) + b^2(v) + M\eta^1(v) + \eta^2(v)\bigr)e(v)^{-\gamma}\, dm(v), \tag{59}$$

$$\hat{I}_2 \equiv \int \bigl(b^1(v) + \eta^1(v)\bigr)e(v)^{-\gamma}\, dm(v), \tag{60}$$

$$I_3 \equiv \int \bigl(1 - e(v) - g(v) + M\eta^1(v) + \eta^2(v)\bigr)^{1-\sigma}\, dm(v), \tag{61}$$

and

$$I_4 \equiv \int \bigl(1 - e(v) - g(v) + M\eta^1(v) + \eta^2(v)\bigr)^{-\sigma}\, dm(v). \tag{62}$$

In terms of these equations, our problem becomes

$$\max_{e(v)} \bigl[\alpha^1 M I_1^{\gamma-1} \hat{I}_2^{1-\gamma} + \alpha^2 \bigl(1 - M I_1^{-1} \hat{I}_2\bigr)^{1-\gamma}\bigr] \frac{I_1}{1-\gamma} + \alpha^2 \xi \frac{I_3}{1-\sigma} \tag{63}$$

subject to

$$\log \xi + \gamma \log\bigl(1 - M I_1^{-1} \hat{I}_2\bigr) + \log(I_4 - I_3) - \log(I_1 - I_2) = 0. \tag{64}$$

Note that I_3 has a monotone effect on the objective function and that I_4 enters in the constraint but not in the objective function.

Marco Bassetto: bassetto@nber.org

DOI: 10.3982/QE362

If we find a perturbation to a given trajectory $\hat{e}$ that changes I_3, but not I_1, I_2, $\hat{I}_2$, or $I_3 - I_4$, we can improve on $\hat{e}$ while keeping the constraint holding. Therefore, at an optimum this cannot happen.

We will proceed as follows:

(i) Lemma 1 first restricts the cases in which the optimal solution $\hat{e}$ may lie on a boundary, that is, at 0 or at $1-g$.

(ii) Under the special case of constant and deterministic government spending and coupon payments, Lemma 2 shows that $\hat{e}$ can take at most three values, except on a set of m-measure 0.

(iii) Under the special case of Lemma 2, Lemma 3 strengthens the result and proves that $\hat{e}$ may take at most two values only, except on a set of m-measure 0.

(iv) We generalize the proof by removing the assumption that led us to study the special case. As we will see, the proof of the general case will work by "conditioning" on the values of $(g_t, b_t^1, b_t^2, \eta_t^1, \eta_t^2)$ and reducing the problem to our special case. This step will not be trivial but will not contain any further insights.

D.1 *Statement and proof of Lemma 1*

Lemma 1. *Assume that Condition* 1 *holds. Then the optimal choice* $\hat{e}$ *for maximizing* (63) *subject to* (64) *may not be equal to* 0, *except on sets of* m*-measure* 0, *and it can only be equal to* $1 - g + M\eta^1 + \eta^2$ *if it is equal to that value almost everywhere with respect to the measure* m.

Proof. We first consider the boundary $\hat{e}(v) = 0$. If $\gamma \geq 1$, choosing $\hat{e}(v) = 0$ with positive measure leads both consumers to infinite negative utility; the government will never pick such a policy whenever an alternative policy is available.

Consider now the case $\gamma < 1$, $(g_t, \eta_t^2 + b_t^2, \eta_t^1 + b_t^1) \neq \mathbf{0}$. Since $g < 1 + M\eta^1 + \eta^2$ by our assumptions that a competitive equilibrium exists, equation (51) implies $x^2(v) > 0$. The leisure of type-2 agents can be positive while their consumption is 0 only if the tax rate is 100% in the state we are considering, that is, $\tau(v) = 1$. In this case, these agents will not work at all, which implies $x^2(v) = 1$: this can be consistent with market clearing only if $g(v) = M\eta^1(v) + \eta^2(v)$. Since the price of the goods in the states with no consumption is infinite, the budget constraints of the agents can hold only if $\eta^2(v) + b^2(v) = 0$ and $\eta^1(v) + b^1(v) = 0$. These requirements violate Condition 1. We thus proved that $e > 0$ except at most on sets of m-measure 0.

We now look at the consequences of $e(v) = 1 - g(v) + M\eta^1(v) + \eta^2(v)$. From (52), this can happen in two cases: either $k^2 = 0$ or $\tau(v) = -\infty$. The latter case is easily shown to be incompatible with the government budget constraint. If $k^2 = 0$, it follows that $c^2(v) = 0$ a.s. and, hence, from (31), $x^2(v) = 0$ a.s. as well. We thus obtain $e(v) = 1 - g(v) + M\eta^1(v) + \eta^2(v)$ a.s. from equation (51). □

D.2 *Statement and proof of Lemma 2*

LEMMA 2. *Assume that* $(g_t, b_t^1, b_t^2, \eta_t^1, \eta_t^2)$ *take a single value almost everywhere with respect to the measure m. Assume that Condition 1 holds. Then the optimal choice* $\hat{e}$ *for maximizing* (63) *subject to* (64) *may take at most three values, except on a set of m-measure* 0.

Lemma 2 corresponds to the case of no uncertainty, constant government spending, and constant coupon payments among all the agents in the economy.

PROOF OF LEMMA 2. We reason by contradiction. Let $\hat{e}$ be the optimal choice by the government. We ruled out that $\hat{e}(v) = 0$ or $\hat{e}(v) = 1 - g + M\eta^1 + \eta^2$ with positive probability, unless $\hat{e}(v) = 1 - g + M\eta^1 + \eta^2$ a.s., in which case our statement holds.[35] Therefore, if $\hat{e}$ takes more than three values, we can find an open set $S \subset (0, 1 - g + M\eta^1 + \eta^2)$ such that $\hat{e}$ takes more than three values in S.[36] Let $V \equiv \hat{e}^{-1}(S)$ be the set of realizations of v such that $e(v)$ falls into S. We wish to prove that there exists a function e that satisfies the constraint (64) and leads to a higher value for the objective. We restrict our search to the space

$$\mathcal{S} \equiv \left\{e : e \text{ is } m\text{-measurable} \wedge e(v) = \hat{e}(v)\ \forall v \in [0,1] \setminus V \wedge e(v) \in S\ \forall v \in V\right\}. \tag{65}$$

It is easy to see that $\hat{e} \in \mathcal{S}$.

The space $\mathcal{S}$ allows perturbations of $\hat{e}$ only in the range where the function lies in S. The reason for this is to be sure that a Fréchet differential is properly defined.

Note first that if $e \in \mathcal{S}$, then its restriction to V $(\hat{e}|_V)$ belongs to $L_m^1(V)$, which is a Banach space. Furthermore, the space of all the restrictions to V of functions in $\mathcal{S}$ is an open subset of $L_m^1(V)$. Since all the perturbations we consider coincide outside of V by our construction of $\mathcal{S}$, we only consider their restriction on V.

We can treat I_1, I_2, $\hat{I}_2$, I_3, and I_4 as functions of $e|_V$. It is more convenient to replace I_2 and $\hat{I}_2$ in our analysis with

$$\tilde{I}_2 \equiv \int e(v)^{-\gamma}\, dm(v). \tag{66}$$

In the case we are considering here, we have $I_2 = (Mb^1 + b^2 + M\eta^1(v) + \eta^2(v))\tilde{I}_2$ and $\hat{I}_2 = (b^1 + \eta^1)\tilde{I}_2$: both I_2 and $\hat{I}_2$ are simply proportional to $\tilde{I}_2$, so that a perturbation that does not affect the latter integral will not affect the two former integrals either.

Let thus $I \equiv (I_1, \tilde{I}_2, I_3, I_4) : \mathcal{S} \to \mathbb{R}^4$. I is Fréchet differentiable and its Fréchet differential is given by

$$\delta I(e|_V; h) = \begin{bmatrix} (1-\gamma)\int_V h(v)e(v)^{-\gamma}\, dm(v) \\ -\gamma \int_V h(v)e(v)^{-\gamma-1}\, dm(v) \\ (1-\sigma)\int_V h(v)(1 - e(v) - g + M\eta^1 + \eta^2)^{-\sigma}\, dm(v) \\ -\sigma \int_V h(v)(1 - e(v) - g + M\eta^1 + \eta^2)^{-\sigma-1}\, dm(v) \end{bmatrix}. \tag{67}$$

[35] For simplicity of notation, we can here drop the dependence of g, b^1, b^2, and η^1 on v, since they are constant functions almost everywhere with respect to the measure m.

[36] Note that we require S to be a strict subset of $(0, 1 - g + M\eta^1 + \eta^2)$. This is convenient to ensure that all our integrals will be properly defined.

We know that if $\hat{e}$ is a regular point for the mapping I, then we can find a perturbation that will leave I_1, I_2, and $I_3 - I_4$ unchanged while increasing or decreasing I_3.[37] This would imply that it is possible to improve upon the choice of $\hat{e}$ and, therefore, $\hat{e}$ would not be optimal.

We, therefore, need to show that $\hat{e}$ is a regular point for the mapping I whenever it takes more than three values in V with positive measure m. $\hat{e}$ will be a regular point for I whenever its Fréchet differential is onto $\mathbb{R}^4$.

Since the function h is an arbitrary function in $L^1_m(V)$, $\delta I(e|_V; h)$ will not be onto $\mathbb{R}^4$ if and only if there is a nonzero vector $a \equiv (a_1, a_2, -a_3, -a_4)$ such that

$$a \cdot \begin{bmatrix} e(v)^{-\gamma} \\ e(v)^{-\gamma-1} \\ (1-e(v)-g+M\eta^1+\eta^2)^{-\sigma} \\ (1-e(v)-g+M\eta^1+\eta^2)^{-\sigma-1} \end{bmatrix} = 0 \tag{68}$$

for all $v \in V$, except at most a set of m-measure 0.

The remainder of the proof of Lemma 2 shows that equation (68) can never hold in more than three points. To do this, we define

$$f_1(y) \equiv a_1 y^{-\gamma} + a_2 y^{-\gamma-1}, \tag{69}$$

$$f_2(y) \equiv a_3\big(1-y-g+M\eta^1+\eta^2\big)^{-\sigma} + a_4\big(1-y-g+M\eta^1+\eta^2\big)^{-\sigma-1}, \tag{70}$$

and we look for the maximum number of intersections between f_1 and f_2 in $(0, 1-g+M\eta^1+\eta^2)$. By enumerating and studying each possible sign that each component of a can take, it is possible to show that in no case can there be more than three intersections between f_1 and f_2. Note that by linear homogeneity, we can restrict our attention to $a_1 = 1$ or $a_1 = 0$, $a_2 = 1$. I only present here the analysis of the most complicated case, that is, $a_1 = 1$, $a_2 < 0$, $a_3 > 0$, $a_4 < 0$. The other cases are available upon request.[38]

In this case, f_1 is negative[39] for $y < -a_2$, strictly increasing for $y < -\frac{a_2(\gamma+1)}{\gamma}$, strictly concave for $y < -\frac{a_2(\gamma+2)}{\gamma}$, and has a strictly positive third derivative for $y < -\frac{a_2(\gamma+3)}{\gamma}$.

f_2 is strictly positive for $y < 1-g+M\eta^1+\eta^2+\frac{a_4}{a_3}$, strictly increasing for $y < 1-g+M\eta^1+\eta^2+\frac{a_4(\sigma+1)}{a_3\sigma}$, strictly convex for $y < 1-g+M\eta^1+\eta^2+\frac{a_4(\sigma+2)}{a_3\sigma}$, and has a strictly positive third derivative for $y < 1-g+M\eta^1+\eta^2+\frac{a_4(\sigma+3)}{a_3\sigma}$.

We will prove that $f_1' - f_2'$ has at most two roots over $(0, 1-g+M\eta^1+\eta^2)$. This is enough to establish that $f_1 - f_2$ has at most three roots. We distinguish seven subcases.

1. $-\frac{a_2(\gamma+1)}{\gamma} \le 1-g+M\eta^1+\eta^2+\frac{a_4(\sigma+2)}{a_3\sigma}$. In all of these subcases, $f_1' - f_2'$ has exactly one root in the interval $(0, -\frac{a_2(\gamma+1)}{\gamma}]$: $f_1' - f_2'$ is strictly decreasing; it converges to $+\infty$ as $y \to 0$

[37] This is an application of Theorem 1 in Section 9.2 of Luenberger (1969).

[38] All the other cases are considerably simpler and most of them are trivial. In particular, this is the only case for which we need to study derivatives of up to the third order!

[39] The complete statement would say that f_1 is strictly negative for $y < -a_2$, 0 for $y = -a_2$, and positive for $y > -a_2$. In this statement and all the following ones, we will leave the equality and the other side of the inequality implicit. This is just for brevity.

and it is strictly negative in $-\frac{a_2(\gamma+1)}{\gamma}$. Furthermore, $f_1' - f_2'$ has no roots $(-\frac{a_2(\gamma+1)}{\gamma}, 1-g+M\eta^1+\eta^2+\frac{a_4(\sigma+1)}{a_3\sigma}]$, since it is strictly negative in this interval.

1a. $-\frac{a_2(\gamma+1)}{\gamma} < -\frac{a_2(\gamma+2)}{\gamma} \leq 1-g+M\eta^1+\eta^2+\frac{a_4(\sigma+1)}{a_3\sigma}$. In this case, $f_1' - f_2'$ is strictly increasing on $(1-g+M\eta^1+\eta^2+\frac{a_4(\sigma+1)}{a_3\sigma}, 1-g+M\eta^1+\eta^2)$. It is strictly negative at the lower bound and tends to $+\infty$ at the upper bound; we, therefore, have exactly one intersection. In subcase 1a, we, therefore, have exactly two intersections between f_1' and f_2' in $(0, 1-g+M\eta^1+\eta^2)$.

1b. $-\frac{a_2(\gamma+1)}{\gamma} < 1-g+M\eta^1+\eta^2+\frac{a_4(\sigma+1)}{a_3\sigma} < -\frac{a_2(\gamma+2)}{\gamma} < 1-g+M\eta^1+\eta^2$. Let us first consider the interval $(1-g+M\eta^1+\eta^2+\frac{a_4(\sigma+1)}{a_3\sigma}, -\frac{a_2(\gamma+2)}{\gamma}]$. In this interval, f_1' is strictly convex, whereas f_2' is strictly concave; it follows that $f_1' - f_2'$ is strictly convex. Since $f_1' - f_2'$ is strictly negative at the lower bound of the interval, it can have either zero or one roots in the interval, depending on the sign it takes at the upper bound. In the interval $(-\frac{a_2(\gamma+2)}{\gamma}, 1-g+M\eta^1+\eta^2)$, $f_1' - f_2'$ is strictly increasing and its limit at $1-g+M\eta^1+\eta^2$ is $+\infty$; if it is nonnegative at the lower bound, this implies that there was exactly one root in $(1-g+M\eta^1+\eta^2+\frac{a_4(\sigma+1)}{a_3\sigma}, -\frac{a_2(\gamma+2)}{\gamma}]$ and there is no root in $(-\frac{a_2(\gamma+2)}{\gamma}, 1-g+M\eta^1+\eta^2)$; if it is strictly negative at the lower bound, then there was no root in $(1-g+M\eta^1+\eta^2+\frac{a_4(\sigma+1)}{a_3\sigma}, -\frac{a_2(\gamma+2)}{\gamma}]$ and there is exactly one root in $(-\frac{a_2(\gamma+2)}{\gamma}, 1-g+M\eta^1+\eta^2)$. It follows that in subcase 1b, we have exactly two intersections between f_1' and f_2' on $(0, 1-g+M\eta^1+\eta^2)$.

1c. $-\frac{a_2(\gamma+1)}{\gamma} < 1-g+M\eta^1+\eta^2+\frac{a_4(\sigma+1)}{a_3\sigma} < 1-g+M\eta^1+\eta^2 \leq -\frac{a_2(\gamma+2)}{\gamma}$. In this case, $f_1' - f_2'$ is convex over the interval $(1-g+M\eta^1+\eta^2+\frac{a_4(\sigma+1)}{a_3\sigma}, 1-g+M\eta^1+\eta^2)$; it is strictly negative at the lower bound and it converges to $+\infty$ at the upper bound, so that it has exactly one intersection in the considered interval. In subcase 1c, we thus have exactly two intersections between f_1' and f_2' on $(0, 1-g+M\eta^1+\eta^2)$.

2. $1-g+M\eta^1+\eta^2+\frac{a_4(\sigma+2)}{a_3\sigma} < -\frac{a_2(\gamma+1)}{\gamma}$.

2a. $0 < 1-g+M\eta^1+\eta^2+\frac{a_4(\sigma+2)}{a_3\sigma} < -\frac{a_2(\gamma+2)}{\gamma} < 1-g+M\eta^1+\eta^2$. On $(0, 1-g+M\eta^1+\eta^2+\frac{a_4(\sigma+2)}{a_3\sigma}]$, $f_1' - f_2'$ is strictly decreasing; its limit at the lower bound is $+\infty$. There are no roots if $f_1' - f_2'$ is positive at the upper bound, and exactly one root if $f_1' - f_2'$ is nonpositive at the upper bound. On $(1-g+M\eta^1+\eta^2+\frac{a_4(\sigma+2)}{a_3\sigma}, -\frac{a_2(\gamma+2)}{\gamma}]$, $f_1' - f_2'$ is strictly convex. If $f_1' - f_2'$ is positive at the lower bound, it can have zero, one, or two roots in the interval $(1-g+M\eta^1+\eta^2+\frac{a_4(\sigma+2)}{a_3\sigma}, -\frac{a_2(\gamma+2)}{\gamma}]$ and, thus, the same number of roots in $(0, -\frac{a_2(\gamma+2)}{\gamma}]$. If $f_1' - f_2'$ is nonpositive at $1-g+M\eta^1+\eta^2+\frac{a_4(\sigma+2)}{a_3\sigma}$, it can have zero or one roots in $(1-g+M\eta^1+\eta^2+\frac{a_4(\sigma+2)}{a_3\sigma}, -\frac{a_2(\gamma+2)}{\gamma}]$ and, thus, it will have either one or two roots in $(0, -\frac{a_2(\gamma+2)}{\gamma}]$. On $(-\frac{a_2(\gamma+2)}{\gamma}, 1-g+M\eta^1+\eta^2)$, $f_1' - f_2'$ is strictly increasing and its limit at the upper bound is $+\infty$. If $f_1' - f_2'$ has an even number of roots in $(0, -\frac{a_2(\gamma+2)}{\gamma}]$, then it is nonnegative at $-\frac{a_2(\gamma+2)}{\gamma}$ and, thus, there are no roots in $(-\frac{a_2(\gamma+2)}{\gamma}, 1-g+M\eta^1+\eta^2)$; if it has an odd number of roots in $(0, -\frac{a_2(\gamma+2)}{\gamma}]$, then it is negative at $-\frac{a_2(\gamma+2)}{\gamma}$ and there is exactly one root in $(-\frac{a_2(\gamma+2)}{\gamma}, 1-g+M\eta^1+\eta^2)$. It follows that in subcase 2a, $f_1' - f_2'$ has either zero or two roots over $(0, 1-g+M\eta^1+\eta^2)$.

2b. $1-g+M\eta^1+\eta^2+\frac{a_4(\sigma+2)}{a_3\sigma} \leq 0 < -\frac{a_2(\gamma+2)}{\gamma} < 1-g+M\eta^1+\eta^2$. On $(0, -\frac{a_2(\gamma+2)}{\gamma}]$, $f_1' - f_2'$ is strictly convex and its limit at 0 is $+\infty$. Therefore, it can have zero, one, or two roots in this interval. On $(-\frac{a_2(\gamma+2)}{\gamma}, 1-g+M\eta^1+\eta^2)$, $f_1'-f_2'$ is strictly increasing and its limit at the upper bound is $+\infty$. If $f_1'-f_2'$ has an even number of roots in $(0, -\frac{a_2(\gamma+2)}{\gamma}]$, then it is nonnegative at $-\frac{a_2(\gamma+2)}{\gamma}$ and, thus, there are no roots in $(-\frac{a_2(\gamma+2)}{\gamma}, 1-g+M\eta^1+\eta^2)$; if it has an odd number of roots in $(0, -\frac{a_2(\gamma+2)}{\gamma}]$, then it is negative at $-\frac{a_2(\gamma+2)}{\gamma}$ and there is exactly one root in $(-\frac{a_2(\gamma+2)}{\gamma}, 1-g+M\eta^1+\eta^2)$. Therefore, in subcase 2b, $f_1'-f_2'$ has either zero or two roots in $(0, 1-g+M\eta^1+\eta^2)$.

2c. $0 < 1-g+M\eta^1+\eta^2+\frac{a_4(\sigma+2)}{a_3\sigma} < 1-g+M\eta^1+\eta^2 \leq -\frac{a_2(\gamma+2)}{\gamma}$. On $(0, 1-g+M\eta^1+\eta^2+\frac{a_4(\sigma+2)}{a_3\sigma}]$, $f_1'-f_2'$ is strictly decreasing; its limit at the lower bound is $+\infty$. There are no roots if $f_1'-f_2'$ is positive at the upper bound, and exactly one root if $f_1'-f_2'$ is nonpositive at the upper bound. On $(1-g+M\eta^1+\eta^2+\frac{a_4(\sigma+2)}{a_3\sigma}, 1-g+M\eta^1+\eta^2)$, $f_1'-f_2'$ is strictly convex and its limit at $1-g+M\eta^1+\eta^2$ is $+\infty$. If $f_1'-f_2'$ is positive at $1-g+M\eta^1+\eta^2+\frac{a_4(\sigma+2)}{a_3\sigma}$, then there can be either zero or two roots in this interval. If $f_1'-f_2'$ is nonnegative at $1-g+M\eta^1+\eta^2+\frac{a_4(\sigma+2)}{a_3\sigma}$, since it is strictly decreasing in that point, it follows that it will have exactly one root in $(1-g+M\eta^1+\eta^2+\frac{a_4(\sigma+2)}{a_3\sigma}, 1-g+M\eta^1+\eta^2)$. We thus have that in subcase 2c, there can be either zero or two roots for $f_1'-f_2'$ in $(0, 1-g+M\eta^1+\eta^2)$.

2d. $1-g+M\eta^1+\eta^2+\frac{a_4(\sigma+2)}{a_3\sigma} \leq 0 < 1-g+M\eta^1+\eta^2 \leq -\frac{a_2(\gamma+2)}{\gamma}$. In this case, $f_1'-f_2'$ is convex over the whole interval $(0, 1-g+M\eta^1+\eta^2)$; furthermore, its limits at both bounds are $+\infty$. It follows that it can have either zero or two roots in the interval. □

D.3 *Statement and proof of Lemma* 3

LEMMA 3. *Assume that* $(g_t, b_t^1, b_t^2, \eta_t^1, \eta_t^2)$ *take a single value almost everywhere with respect to the measure* m. *Assume that Condition* 1 *holds. Then the optimal choice* $\hat{e}$ *for maximizing* (63) *subject to* (64) *may take at most two values, except on a set of* m-*measure* 0.

Lemma 3 still considers the case of no uncertainty, constant government spending, and constant coupon payments among all the agents in the economy. It starts from the result of Lemma 2 and strengthens it by considering a different class of perturbations.

PROOF OF LEMMA 3. From Lemma 2, we know that the optimal choice $\hat{e}$ is a step function with at most three values, aside from sets of m-measure 0; from Lemma 1, we know that each of the three values lies in $(0, 1-g+M\eta^1+\eta^2)$, unless $\hat{e}$ is the constant $1-g+M\eta^1+\eta^2$.

We reason again by contradiction. Suppose $\hat{e}$ takes three values, each with positive m-measure; let $e_1 < e_2 < e_3$ be the three values. The integrals I can then be rewritten as

$$I_1 = m_1 e_1^{1-\gamma} + m_2 e_2^{1-\gamma} + \left(\frac{1}{1-\beta} - m_1 - m_2\right) e_3^{1-\gamma}, \tag{71}$$

$$\tilde{I}_2 = m_1 e_1^{-\gamma} + m_2 e_2^{-\gamma} + \left(\frac{1}{1-\beta} - m_1 - m_2\right) e_3^{-\gamma}, \tag{72}$$

$$\begin{aligned} I_3 = {} & m_1\left(1-g+M\eta^1+\eta^2-e_1\right)^{1-\sigma} + m_2\left(1-g+M\eta^1+\eta^2-e_2\right)^{1-\sigma} \\ & + \left(\frac{1}{1-\beta} - m_1 - m_2\right)\left(1-g+M\eta^1+\eta^2-e_3\right)^{1-\sigma}, \end{aligned} \tag{73}$$

$$\begin{aligned} I_4 = {} & m_1\left(1-g+M\eta^1+\eta^2-e_1\right)^{-\sigma} + m_2\left(1-g+M\eta^1+\eta^2-e_2\right)^{-\sigma} \\ & + \left(\frac{1}{1-\beta} - m_1 - m_2\right)\left(1-g+M\eta^1+\eta^2-e_3\right)^{-\sigma}, \end{aligned} \tag{74}$$

where $m_i \equiv m(\{v : e(v) = e_i\})$, $i = 1, 2$.

As we already observed, $\hat{e}$ cannot be optimal if we can perturb I_3 in either direction while holding I_1, $\tilde{I}_2$, and $I_3 - I_4$ constant. For this proof, we treat I as a function of $(e_1, e_2, e_3, m_1, m_2)$. If $\hat{e}$ takes all three values with positive measure, we have $m_i > 0$, $i = 1, 2$, and $m_1 + m_2 < \frac{1}{1-\beta}$. In this case, therefore, I is now a mapping from $\mathbb{R}^5$ to $\mathbb{R}^4$; given our previous observations, the mapping is well defined and differentiable in an open neighborhood of $(e_1, e_2, e_3, m_1, m_2)$. By the same theorem we applied in Lemma 2, $\hat{e}$ cannot be optimal if $(e_1, e_2, e_3, m_1, m_2)$ is a regular point of the mapping I, that is, if the differential of I as a function of $(e_1, e_2, e_3, m_1, m_2)$ is onto $\mathbb{R}^4$. We now prove that $(e_1, e_2, e_3, m_1, m_2)$ is indeed a regular point of I when all three points are distinct and all measures are strictly positive. To do this, we will just perturb (e_1, e_2, e_3, m_2) while we will hold m_1 fixed: we will show that the differential with respect to just the four elements already spans $\mathbb{R}^4$. The Jacobian of the mapping I is given by[40]

$$J = \begin{bmatrix} e_1^{-\gamma} & e_1^{-\gamma-1} & (1-e_1-g+M\eta^1+\eta^2)^{-\sigma} & (1-e_1-g+M\eta^1+\eta^2)^{-\sigma-1} \\ e_2^{-\gamma} & e_2^{-\gamma-1} & (1-e_2-g+M\eta^1+\eta^2)^{-\sigma} & (1-e_2-g+M\eta^1+\eta^2)^{-\sigma-1} \\ e_3^{-\gamma} & e_3^{-\gamma-1} & (1-e_3-g+M\eta^1+\eta^2)^{-\sigma} & (1-e_3-g+M\eta^1+\eta^2)^{-\sigma-1} \\ \frac{e_2^{1-\gamma}-e_3^{1-\gamma}}{1-\gamma} & -\frac{e_2^{-\gamma}-e_3^{-\gamma}}{\gamma} & \frac{(1-e_2-g+M\eta^1+\eta^2)^{1-\sigma}-(1-e_3-g+M\eta^1+\eta^2)^{1-\sigma}}{1-\sigma} & -\frac{(1-e_2-g+M\eta^1+\eta^2)^{-\sigma}-(1-e_3-g+M\eta^1+\eta^2)^{-\sigma}}{\sigma} \end{bmatrix}, \tag{75}$$

[40]For convenience, we scaled the columns by $\frac{1}{1-\gamma}$, $-\frac{1}{\gamma}$, $\frac{1}{1-\sigma}$, and $-\frac{1}{\sigma}$, respectively. This does not alter the rank of the Jacobian and it allows us to have a shorter expression.

which can be rewritten as

$$J = \begin{bmatrix} e_1^{-\gamma} & e_1^{-\gamma-1} & (1-e_1-g+M\eta^1+\eta^2)^{-\sigma} \\ e_2^{-\gamma} & e_2^{-\gamma-1} & (1-e_2-g+M\eta^1+\eta^2)^{-\sigma} \\ e_3^{-\gamma} & e_3^{-\gamma-1} & (1-e_3-g+M\eta^1+\eta^2)^{-\sigma} \\ \int_{e_2}^{e_3} y^{-\gamma}\,dy & \int_{e_2}^{e_3} y^{-\gamma-1}\,dy & \int_{e_2}^{e_3}(1-y-g+M\eta^1+\eta^2)^{-\sigma}\,dy \end{bmatrix} \tag{76}$$
$$\begin{bmatrix} (1-e_1-g+M\eta^1+\eta^2)^{-\sigma-1} \\ (1-e_2-g+M\eta^1+\eta^2)^{-\sigma-1} \\ (1-e_3-g+M\eta^1+\eta^2)^{-\sigma-1} \\ \int_{e_2}^{e_3}(1-y-g+M\eta^1+\eta^2)^{-\sigma-1}\,dy \end{bmatrix}.$$

The Jacobian J can only be singular if there exists a nonzero vector (a_1, a_2, a_3, a_4) such that

$$\begin{aligned} &a_1 e_i^{-\gamma} + a_2 e_i^{-\gamma-1} + a_3\big(1-e_i-g+M\eta^1+\eta^2\big)^{-\sigma} \\ &\quad + a_4\big(1-e_i-g+M\eta^1+\eta^2\big)^{-\sigma-1} = 0, \quad i=1,2,3, \end{aligned} \tag{77}$$

and

$$\begin{aligned} &\int_{e_2}^{e_3} \big(a_1 y^{-\gamma} + a_2 y^{-\gamma-1} + a_3\big(1-y-g+M\eta^1+\eta^2\big)^{-\sigma} \\ &\quad + a_4\big(1-y-g+M\eta^1+\eta^2\big)^{-\sigma-1}\big)\,dy = 0. \end{aligned} \tag{78}$$

From Lemma 2, we know that the function that we are integrating in (78) can have at most three zeros in $(0, 1-g+M\eta^1+\eta^2)$. By (77), the three zeros are e_1, e_2, and e_3, so that the function is never zero in any point of (e_2, e_3); since it is a continuous function, it is either always strictly positive or always strictly negative. It follows that its integral cannot be zero; therefore, J is of full rank and $\hat{e}$ cannot be an optimal choice. □

We are now ready to prove the main theorem.

D.4 *Proof of the main body of Theorem* 1

Note that maximizing (56) subject to (53) can be rewritten as maximizing (63) subject to (64), given the definitions of our integrals. As in Lemmas 2 and 3, our proof proceeds by using the fact that we can improve upon $\hat{e}$ if we can find a perturbation that can vary I_3 in either direction while leaving I_1, I_2, $\hat{I}_2$, and $I_3 - I_4$ unchanged. Recalling the definition of $\tilde{v}$ in (57), we can use Fubini's theorem and rewrite the integrals as

$$I_1 \equiv \int\int_0^1 e(\tilde{v},h)^{1-\gamma}\,dh\,dm\big(\tilde{v},[0,1]\big), \tag{79}$$

$$I_2 \equiv \int\int_0^1 \big(Mb^1(\tilde{v}) + b^2(\tilde{v}) + M\eta^1(\tilde{v}) + \eta^2(\tilde{v})\big)e(\tilde{v},h)^{-\gamma}\,dh\,dm\big(\tilde{v},[0,1]\big), \tag{80}$$

$$\hat{I}_2 \equiv \int\int_0^1 \big(b^1(\tilde{v}) + \eta^1(\tilde{v})\big)e(\tilde{v},h)^{-\gamma}\,dh\,dm\big(\tilde{v},[0,1]\big), \tag{81}$$

$$I_3 \equiv \int \int_0^1 \left(1 - e(\tilde{v}, h) - g(\tilde{v}) + M\eta^1(\tilde{v}) + \eta^2(\tilde{v})\right)^{1-\sigma} dh\, dm\left(\tilde{v}, [0,1]\right), \tag{82}$$

$$I_4 \equiv \int \int_0^1 \left(1 - e(\tilde{v}, h) - g(\tilde{v}) + M\eta^1(\tilde{v}) + \eta^2(\tilde{v})\right)^{-\sigma} dh\, dm\left(\tilde{v}, [0,1]\right). \tag{83}$$

Consider now the inner integrals. In these integrals we are conditioning on $\tilde{v}$ and integrating with respect to h alone. By the same proof as Lemmas 2 and 3, $\hat{e}(\tilde{v}, h)$ must take at most two values as a function of h for each $\tilde{v}$, except at most in sets of Lebesgue measure 0,[41] for otherwise we can vary the inner integral in I_3 while holding the inner integrals in I_1, I_2, $\hat{I}_2$, and $I_3 - I_4$ fixed. Of course, changes in the inner integrals will be reflected in changes in the whole integrals only if they take place on sets that have positive m-measure in the outside integration: therefore, $\hat{e}(\tilde{v}, h)$ can take more than two values as a function of h for any given $\tilde{v}$ on sets of m-measure 0, but this cannot happen on sets of positive m-measure. □

Appendix E: Uniform commodity taxation: A formal analysis

This appendix contains a formal treatment of the relationship between the results in the main paper and the conditions under which uniform commodity taxation holds. In this appendix, we adopt a notation that allows us to easily compare the results in the paper with what has already been established in a static framework, in particular, by Atkinson and Stiglitz (1972) and Atkinson and Stiglitz (1976).

We indicate by c^i, $i = 1, 2$, the vector of consumption goods consumed by type-i agents. x is the vector of leisure consumed by the taxpayers (agents of type 2). The preferences of the rentiers are described by

$$V^1\left(\Gamma^1(c^1)\right) \tag{84}$$

and those of the taxpayers by

$$V^2\left(\Gamma^2(c^2), \Theta(x)\right), \tag{85}$$

where Γ^i, $i = 1, 2$, and Θ are linearly homogeneous functions, and all the functions are assumed to be twice continuously differentiable. Equations (84) and (85) capture two of the features that are relevant for our purposes: that preferences are separable between leisure and the consumption goods, and that the subutilities are homothetic. These assumptions are satisfied by the preferences (14) and (15) that we assumed in Section 3.

To keep notation simple, we will let c^i and x be finite-dimensional vectors; c^i_j will denote the jth component of c^i and x_j will denote the jth component of x. All the results continue to hold if we switch to the appropriate notation in an infinite-dimensional space.

The technology of the economy is characterized by

$$F\left(N^1 c^1 + c^2 + g, x\right) \le 0. \tag{86}$$

[41] Note that h is distributed uniformly, so its measure is the Lebesgue measure.

We assume the technology exhibits constant returns to scale. To stay closer to Atkinson and Stiglitz (A-S), we first assume F to be twice continuously differentiable, with a strictly positive gradient. This assumption implies that any good (or leisure) can be transformed into another good (leisure), which is violated by our problem; we, therefore, will later amend this hypothesis and look at the implications of doing so.

Atkinson and Stiglitz work mostly with a small open economy (or a linear technology) in which producer prices are given, although their results are more general; in their case, the function F could be written as

$$F\left(N^1 c^1 + c^2 + g, x\right) = \sum_j q_j^* \left(N^1 c_j^1 + c_j^2 + g_j\right) + \sum_j w_j^* (x_j - 1), \tag{87}$$

where q_j^* are the international prices of the different consumption goods and w_j^* are the international wages for the various types of leisure.

Let w be the vector of wages corresponding to the different types of leisure and let q be the vector of producer prices of the consumption goods. If we normalize to 1 the wage rate of time of the first type, profit maximization on the firms' part requires

$$q_j = -\frac{F_{c_j}}{F_{x_1}} \tag{88}$$

and

$$w_j = \frac{F_{x_j}}{F_{x_1}}. \tag{89}$$

The budget constraints of the rentiers and the taxpayers can be written as

$$\sum_j p_j \left(c_j^1 - \bar{c}_j^1\right) - T \le 0 \tag{90}$$

and

$$\sum_j p_j \left(c_j^2 - \bar{c}_j^2\right) + \sum_j w_j^a (x_j - 1) - T \le 0, \tag{91}$$

where $\bar{c}^i$ is the vector of the initial endowment of each type, p is the vector of consumer prices, w^a is the vector of after-tax wages, and T is a lump-sum transfer from the government. We already imposed that the taxpayers start with 1 unit of time of each type; since we are free to adjust the function F, this can be viewed simply as a normalization.

In line with A-S, we assume the government can tax the consumption goods (net of the initial endowment) and the labor supply, but it cannot tax any type of leisure. For a general production function F, this is a richer set of instruments than the one we introduced in the paper, where only the labor supply can be taxed. However, we will show later that taxing consumption in addition to the labor supply is redundant for the particular production function that we use in the paper. One tax rate is redundant, so we can set $q_1 = p_1$.

As in the main text, we will work with the primal problem: we will use the first-order conditions of the consumers and the producers to substitute out the prices, and we will look at the Ramsey problem as one of solving for quantities.[42] From the budget constraints and the first-order conditions of the consumers, we obtain the implementability constraints

$$\sum_j \Gamma_j^1 (c_j^1 - \bar{c}_j^1) - \Gamma_1^1 T \leq 0 \tag{92}$$

and

$$V_1^2 \sum_j \Gamma_j^2 (c_j^2 - \bar{c}_j^2) + V_2^2 \sum_j \Theta_j (x_j - 1) - V_1^2 \Gamma_1^2 T \leq 0. \tag{93}$$

In equations (92) and (93) and in what follows, a subscript j to a function refers to the partial derivative with respect to the jth component. We normalized the price of the first consumption good to 1, we multiplied the first equation by Γ_1^1, and multiplied the second equation by $V_1^2 \Gamma_1^2$.

In addition to the implementability constraints, the government faces the following further constraints:

(i) The feasibility constraint, given by equation (86).

(ii) In a competitive equilibrium, the marginal rates of substitution must be the same for all consumers, that is,

$$\Gamma_j^1 \Gamma_k^2 = \Gamma_k^1 \Gamma_j^2 \quad \forall j, k. \tag{94}$$

Because of (94), the implementability constraint of the rentiers can also be written in the following form, which will be more convenient later:

$$V_1^2 \sum_j \Gamma_j^2 (c_j^1 - \bar{c}_j^1) - V_1^2 \Gamma_1^2 T \leq 0. \tag{95}$$

The first-order conditions for the government are

$$\alpha^1 V_1^1 \Gamma_i^1 + \lambda^1 V_1^2 \Gamma_i^2 + \sum_{j>1} \nu_j \left[\Gamma_{1i}^1 \Gamma_j^2 - \Gamma_{ij}^1 \Gamma_1^2\right] = \mu F_{c_i} \quad \forall i, \tag{96}$$

$$\begin{aligned} &\alpha^2 V_1^2 \Gamma_i^2 + \lambda^1 \left\{ V_{11}^2 \Gamma_i^2 \left[\sum_j \Gamma_j^2 (c_j^1 - \bar{c}_j^1) - \Gamma_1^2 T \right] + V_1^2 \left[\sum_j \Gamma_{ij}^2 (c_j^1 - \bar{c}_j^1) - \Gamma_{1i}^2 T \right] \right\} \\ &\quad + \lambda^2 \left\{ V_{11}^2 \Gamma_i^2 \left[\sum_j \Gamma_j^2 (c_j^2 - \bar{c}_j^2) - \Gamma_1^2 T \right] + V_1^2 \left[\sum_j \Gamma_{ij}^2 (c_j^2 - \bar{c}_j^2) - \Gamma_{1i}^2 T \right] \right. \\ &\quad \left. + V_1^2 \Gamma_i^2 + V_{12}^2 \Gamma_i^2 \sum_j \Theta_j (x_j - 1) \right\} + \sum_{j>1} \nu_j \left[\Gamma_1^1 \Gamma_{ij}^2 - \Gamma_j^1 \Gamma_{1i}^2\right] \\ &= \mu F_{c_i} \quad \forall i, \end{aligned} \tag{97}$$

[42] Atkinson and Stiglitz follow the dual approach: they substitute out quantities and solve the problem in terms of prices. The primal approach is easier in our case in which we have an initial endowment of more than one good.

$$\begin{aligned}
&\alpha^2 V_2^2 \Theta_i + \lambda^1 V_{12}^2 \Theta_i \Big[\sum_j \Gamma_j^2 (c_j^1 - \bar{c}_j^1) - \Gamma_1^2 T \Big] \\
&\qquad + \lambda^2 \Big\{ V_{12}^2 \Theta_i \Big[\sum_j \Gamma_j^2 (c_j^2 - \bar{c}_j^2) - \Gamma_1^2 T \Big] + V_{22}^2 \Theta_i \sum_j \Theta_j (x_j - 1) \\
&\qquad + V_2^2 \sum_j \Theta_{ij} (x_j - 1) + V_2^2 \Theta_i \Big\} \\
&\quad = \mu F_{x_i} \quad \forall i,
\end{aligned} \tag{98}$$

and

$$-\lambda^1 V_1^2 \Gamma_1^2 - \lambda^2 V_1^2 \Gamma_1^2 \geq 0 \quad \Longrightarrow \quad \lambda^1 \geq -\lambda^2, T \geq 0, \big(\lambda^1 + \lambda^2\big) T = 0, \tag{99}$$

where λ^1, λ^2, ν, and μ are the Lagrange multipliers associated with the constraints (95), (93), (94), and (86), respectively. Given the Ramsey allocation, the conditions for a competitive equilibrium imply the price system and tax policy

$$p_i = \frac{\Gamma_i^2}{\Gamma_1^2} \quad \forall i, \tag{100}$$

$$w_i^a = \frac{V_2^2 \Theta_i}{V_1^2 \Gamma_1^2} \quad \forall i, \tag{101}$$

$$q_i = \frac{F_{c_i}}{F_{c_1}} \quad \forall i, \tag{102}$$

$$w_i = \frac{F_{x_i}}{F_{c_1}} \quad \forall i, \tag{103}$$

$$\tau_i^c = \frac{p_i}{q_i} - 1 \quad \forall i, \tag{104}$$

and

$$\tau_i^w = 1 - \frac{w_i^a}{w_i} \quad \forall i, \tag{105}$$

where we normalized $p_1 = q_1 = 1$.

We have a uniform commodity tax when $\frac{p_i}{q_i}$ is independent of i or, equivalently, when $\frac{\Gamma_i^2}{F_{c_i}}$ is independent of i.[43] To study what conditions lead to a uniform commodity tax, it is useful to rewrite the first-order conditions as

$$\alpha^1 V_1^1 \frac{\Gamma_i^1}{F_{c_i}} + \lambda^1 V_1^2 \frac{\Gamma_i^2}{F_{c_i}} = \mu - \frac{1}{F_{c_i}} \sum_{j>1} \nu_j \big[\Gamma_{1i}^1 \Gamma_j^2 - \Gamma_{ij}^1 \Gamma_1^2 \big] \quad \forall i, \tag{106}$$

[43] Due to (94), this also implies that $\frac{\Gamma_i^1}{F_{c_i}}$ is independent of i.

$$\begin{aligned}
&\frac{\Gamma_i^2}{F_{c_i}}\Bigg\{\alpha^2 V_1^2 + \lambda^1\Bigg[V_{11}^2\sum_j \Gamma_j^2(c_j^1-\bar{c}_j^1)\Bigg] \\
&\quad + \lambda^2\Bigg[V_{11}^2\sum_j \Gamma_j^2(c_j^2-\bar{c}_j^2) + V_1^2 + V_{12}^2\sum_j \Theta_j(x_j-1)\Bigg]\Bigg\} \\
&= \mu - \frac{1}{F_{c_i}}\Bigg[\lambda^1 V_1^2\sum_j \Gamma_{ij}^2(c_j^1-\bar{c}_j^1) \\
&\quad + \lambda^2 V_1^2\sum_j \Gamma_{ij}^2(c_j^2-\bar{c}_j^2) + \sum_{j>1}\nu_j(\Gamma_1^1\Gamma_{ij}^2 - \Gamma_j^1\Gamma_{1i}^2)\Bigg] \quad \forall i,
\end{aligned} \tag{107}$$

and

$$\begin{aligned}
&\frac{\Theta_i}{F_{x_i}}\Bigg\{\alpha^2 V_2^2 + \lambda^1 V_{12}^2\sum_j \Gamma_j^2(c_j^1-\bar{c}_j^1) \\
&\quad + \lambda^2\Bigg[V_{12}^2\sum_j \Gamma_j^2(c_j^2-\bar{c}_j^2) + V_{22}^2\sum_j \Theta_j(x_j-1) + V_2^2\Bigg]\Bigg\} \\
&= \mu - \frac{\lambda^2 V_2^2}{F_{x_i}}\sum_j \Theta_{ij}(x_j-1) \quad \forall i
\end{aligned} \tag{108}$$

together with (99).

Note that all the terms in T dropped out of equations (106), (107), and (108) because of (99). Lump-sum transfers and taxes affect our problem only through the multipliers λ^1 and λ^2: given such multipliers, the first-order conditions are identical whether T is optimally chosen or is constrained to be 0, as in our main text. While lump-sum taxes and transfers can reduce the incentive to distort prices, they cannot completely offset it unless $\lambda^1 = \lambda^2 = 0$; it is easy to check that in this case the allocation is an unconstrained Pareto optimum, which can only happen if the government does not need to levy distortionary taxes.

We can now focus on the terms that break the optimality of a uniform commodity tax.

(i) The terms $\sum_j \Gamma_{ij}^2 c_j^1$ and $\sum_j \Gamma_{ij}^2 c_j^2$. If Γ^2 is homogeneous of degree 1, then its derivatives are homogeneous of degree 0 and hence $\sum_j \Gamma_{ij}^2 c_j^2 = 0$: this is the key to the result obtained by Atkinson and Stiglitz (1972). However, it is not enough for Γ^2 (or Γ^1) to be homogeneous of degree 1 to reduce the first of the two terms to 0. If Γ^1 and Γ^2 are two different homogeneous functions, then the rentiers and the taxpayers will allocate their spending over the consumption goods in different proportions; the government will then be able to favor a group by taxing more lightly the goods it consumes in a larger proportion, which would lead away from uniform taxation. On the other hand, if Γ^1 and Γ^2 are the same function, then equality of the marginal rates of substitution implies that c^1 is proportional to c^2 and both sums are 0: in this case, both groups allocate their spending in equal proportions on the consumption goods and the government would not favor either by deviating from a uniform tax. The environment of Sections 4 and 5 satisfies the condition $\Gamma^1 = \Gamma^2$, so the deviations from a uniform commodity tax do not arise from this source in our case.

(ii) The terms $\sum_j \Gamma^2_{ij}\bar{c}^1_j$ and $\sum_j \Gamma^2_{ij}\bar{c}^2_j$. Even if Γ^1 and Γ^2 are homogeneous of degree 1 and are the same function, deviations from a uniform tax follow if $\bar{c}^1$ and $\bar{c}^2$ are not proportional to c^2 (and c^1). Since $\bar{c}^1$ is the only source of resources for the rentiers, if it is proportional to c^1, it must be equal to it as well: we are then in a case in which the rentiers do not trade away from their initial endowment. As we observed, the net trade between the two types is the main determinant of the pattern of taxes we derive in this paper. By distorting prices, taxes change the value of the initial endowment at consumer prices; each group gets a positive (negative) income effect from increases in the prices of goods for which it is a net seller (buyer). For this reason, a government that wants to favor the rentiers will use some price distortion even if it can raise revenues from the taxpayers and redistribute them lump sum; the first-order conditions show that it will be optimal to trade off the increased distortions from the necessity of raising additional revenues to rebate lump sum with the price distortions that a nonuniform commodity tax implies.

(iii) The term $\sum_{j>1} \nu_j(\Gamma^1_1\Gamma^2_{ij} - \Gamma^1_j\Gamma^2_{1i})$. This term comes from the constraint that marginal rates of substitution should be equal across consumers (equation (94)). We now show that this constraint is not binding if the previous two sources of deviation from uniform commodity taxation are not present. To see this, assume that the functions Γ^1 and Γ^2 are the same and that $\bar{c}^1$ and $\bar{c}^2$ are proportional to c^1 and c^2. If (94) is not binding, then $\nu = \mathbf{0}$. Under these conditions, equation (107) implies that $\frac{\Gamma^2_i}{F_{c_i}}$ is independent of i and, hence, equation (106) implies the same for $\frac{\Gamma^1_i}{F_{c_i}}$; Γ^1 and Γ^2 are thus proportional to each other and the constraint (94) is satisfied.

Notice that uniform commodity taxation *does not imply* uniform factor taxation. We could easily repeat the same steps to analyze the taxes on factors; since the initial endowment of each factor is 1, homotheticity will not be enough to establish uniform factor taxation unless the labor supply is constant. This is the reason why the labor tax rate is not constant in the representative-agent economy of Lucas and Stokey (1983), even when the separability and homotheticity requirements discussed above hold and when there is no initial government debt.

The previous analysis requires the production function to allow for substitutability of all input and output factors. Our technology is instead described by

$$F_i\left(N^1 c^1_i + c^2_i + g_i, x_i\right) \le 0 \quad \forall i. \tag{109}$$

With this production function, the multiplier μ will be a vector rather than a scalar. In this case, the Ramsey allocation identifies uniquely the consumer prices (through the marginal rates of substitution), but not the producer prices: since there is no substitutability among different goods, firms will not be able to change their production in response to changes in relative prices. Analogously, firms cannot substitute different factors of production and, hence, are not able to react to changes in wages before tax. The firms' profit maximization conditions only link the producer price of a good with the wage before tax in the same period and state of nature. Because of this, the government can then implement the Ramsey allocation by using consumption taxes or labor taxes alone, as we assumed in the text; the only requirement is

$$\frac{1+\tau_j^c}{1-\tau_j^w} = \frac{F_{jx} V_1^2 \Gamma_j^2}{F_{jc} V_2^2 \Theta_j} \quad \forall j. \tag{110}$$

Given equation (110), it is always possible to obtain a constant tax rate on consumption goods or a constant tax rate on all factors of production by an appropriate choice. As an example, in the paper, we normalized $\tau^c = \mathbf{0}$. However, when the sources of deviations from uniform commodity taxes and/or uniform factor taxes are present, it is not possible, in general, to have *both* a uniform commodity tax and a uniform factor tax. For this reason, we obtain different tax rates on labor income even under the assumptions of Section 4.

References

Atkinson, A. and J. E. Stiglitz (1972), "The structure of indirect taxation and economic effciency." *Journal of Public Economics*, 1 (1), 97–119. [9, 13]

Atkinson, A. and J. E. Stiglitz (1976), "The design of tax structure: Direct versus indirect taxation." *Journal of Public Economics*, 6 (1–2), 55–75. [9]

Lucas, R. E., Jr. and N. L. Stokey (1983), "Optimal fiscal and monetary policy in an economy without capital." *Journal of Monetary Economics*, 12, 55–93. [14]

Luenberger, D. G. (1969), *Optimization by Vector Space Methods*. Wiley, New York. [4]

Submitted April, 2013. Final version accepted October, 2013.