# APPENDIX B: TECHNICAL APPENDIX FOR ACEMOGLU-GOLOSOV-TSYVINSKI "**DYNAMIC MIRRLEES TAXATION UNDER POLITICAL ECONOMY CONSTRAINTS**" (NOT FOR PUBLICATION)

## PROPERTIES OF $\mathcal{U}\left(\{C_T, L_T\}_{T=0}^{\infty}\right)$

*Some Technical Results*

We first present some technical results that will be useful in establishing the properties of the functional $\mathcal{U}\left(\{C_t, L_t\}_{t=0}^{\infty}\right)$.

**Definition B1** *Let $X$ and $Z$ be Banach spaces and $G : X \to Z$ be a vector-valued mapping. Suppose that $G$ is continuously (Fréchet) differentiable in the neighborhood of $x_0$ with the derivative denoted by $G'(x_0)$. Then $x_0$ is said to be a* ***regular point*** *of $G$ if $G'(x_0)$ maps $X$ onto $Z$.*

**Lemma B1** *Let $X$ and $Z$ be Banach spaces. Consider the maximization problem of*

$$P(u) = \max_{x \in X} f(x) \tag{B1}$$

*subject to*

$$g_0(x) \leq u \tag{B2}$$

*and*

$$G(x) \leq \mathbf{0} \tag{B3}$$

*where $f : X \to R$ and $g_0 : X \to R$ are real-valued functions and $G : X \to Z$ is a vector-valued mapping and $0$ is the zero of the Banach space $Z$. Suppose that $f$ is concave and $g_0$ is convex, and moreover that the solution at $u = 0$, $x_0$, is a regular point. Let $\mu$ be any multiplier of (B2). Then $\mu$ is a subgradient of $P(0)$.*

*Proof.* This lemma is a direct generalization of Proposition 6.5.8 of Bertsekas, Nedic and Ozdaglar (2003, p. 382) to an infinite dimensional maximization problem. ■

**Theorem B1** *Let $X$ and $Z$ be Banach spaces. Consider the maximization problem of*

$$P(\mathbf{u}) = \max_{x \in X} f(x)$$

*subject to*

$$G(x) \leq \mathbf{0} + \mathbf{u}$$

*where $f : X \to R$ is a real-valued concave function and and $G : X \to Z$ is a convex vector-valued mapping and $0$ is the zero of the vector space $Z$ and $\mathbf{u}$ is a perturbation. Suppose that $x_0$ is a solution to this program. Suppose also that $x_0$ is a regular point of $G$ and that $f$ and $G$ are continuously (Fréchet) differentiable in the neighborhood of $x_0$. Then $P(\mathbf{0})$ is differentiable.*

*Proof.* From Lemma B1, it follows that if there is a unique multiplier, $P$ has a unique subgredient and is thus differentiable. Proposition 4.47 in Bonnans and Shapiro (2000) establishes that under a weaker constraint qualification condition than regularity, this problem has a unique multiplier. ■

**Theorem B2** *Let $X$ be a compact metric space, then the space of probability measures defined on $X$ is a compact metric space with the weak topology.*

*Proof.* See Parthasarathy (1967, p. 45). ■

*Randomizations*

We next introduce randomizations to show concavity and differentiability of $\mathcal{U}\left(\{C_t, L_t\}_{t=0}^{\infty}\right)$. To simplify notation, in this appendix, we suppress dependence on public histories $h^t$. The original maximization problem without randomization is to maximize (3.10) subject to (3.11), (3.12), and (3.14) as stated in Proposition 2. Recall also that $\theta_t \in \Theta$, where $\Theta$ is at finite set (with $N+1$ elements). Therefore $\Theta^t$ for any $t < \infty$ is also a finite set. Consider next the functions $c_t : \Theta^t \to \mathbb{R}_+$ and $l_t : \Theta^t \to \left[0, \bar{l}\right]$. By definition, these functions assign values to a finite number of points in the set $\Theta^t$ for any $t < \infty$, thus can simply be thought of as vectors of $\left(N\left(N+1\right)\right)^t$ dimension. Moreover

$$\int c_t\left(\theta^t\right) dG\left(\theta^t\right) \leq \bar{Y},\ K_{t+1} \leq \bar{Y} \text{ and } x_t \leq \bar{Y}, \tag{B4}$$

where $\bar{Y} = F\left(\bar{Y}, \bar{l}\right) < \infty$. Therefore, $X_t = \left\{c_t\left(\theta^t\right), l\left(\theta^t\right), K_{t+1}, x_t\right\}$ is a vector (of dimension $\left(N\left(N+1\right)\right)^{2t} + 2$). Let $\mathbf{X}_t$ be the set of all such vectors that satisfy the inequalities in (B4), and for $X_t \in \mathbf{X}_t$, let $X_t\left(i\right)$ denote the $i$th component of this vector, and $T_t$ be the dimension of vectors in the set $\mathbf{X}_t$ (i.e., $T_t = \left(N\left(N+1\right)\right)^{2t} + 2$). $\mathbf{X}_t$ is a compact metric space space with the usual Euclidean distance metric, $d_t\left(X_t, X'\right) = \left(\sum_{i=1}^{T_t}\left(X'_t\left(i\right) - X_t\left(i\right)\right)^2\right)^{1/2}$

Let us now construct the product space of the $\mathbf{X}_t$'s

$$\mathbf{X} = \prod_{t=1}^{\infty} \mathbf{X}_t$$

Clearly the sequence $\left\{c_t\left(\theta^t\right), l_t\left(\theta^t\right), K_{t+1}, x_t\right\}_{t=0}^{\infty}$ must belong to $\mathbf{X}$. In fact, it must belong to the subset of $\mathbf{X}$, which satisfies (3.11), (3.12), and (3.14), denoted by $\bar{\mathbf{X}}$.

Now by Tychonoff's theorem (e.g., Dudley, 2002, Theorem 2.2.8), $\mathbf{X}$ is compact in the product topology. Since (3.11), (3.12), and (3.14) are (weak) inequalities, $\bar{\mathbf{X}}$ is a closed subset of $\mathbf{X}$, and therefore it is also compact in the product topology. Moreover, $\mathbf{X}$ with the product topology is meterizable, with the metric

$$d\left(X, X'\right) = \sum_{t=1}^{\infty} \phi^t d_t\left(X_t, X'_t\right) \tag{B5}$$

for some $\phi \in (0,1)$ and $X \equiv \{X_t\}_{t=0}^{\infty} \in \mathbf{X}$. This shows that $\mathbf{X}$ endowed with the product topology is a metric space, and so is $\bar{\mathbf{X}}$.

From Theorem B2, the set of probability measures defined over a compact metric space is compact in the weak topology. This establishes that the set of probability measures $\mathcal{P}^{\infty}$ defined over $\bar{\mathbf{X}}$ is compact in the weak topology.

We are concerned not with all probability measures, but those that condition at $t$ on information revealed up to $t$. Let $\mathcal{C} = \{(c,l) \in \mathbb{R}^2 : 0 \leq c \leq \bar{c},\ 0 \leq l \leq \bar{l}\}$ be the set of possible consumption-labor allocations for agents, so that $\mathcal{P}^{\infty}$ defined above is the set of all probability measures over $\mathcal{C}^{\infty}$. Now, for each $t \in \mathbb{N}$ and $\theta^{t-1} \in \Theta^{t-1}$, let $\mathcal{P}\left[\theta^{t-1}\right]$ be the space of $N+1$-tuples of probability measures on Borel subsets of $\mathcal{C}$ for an individual with history of reports $\theta^{t-1}$. Thus each element $\zeta\left(\cdot \mid \theta^{t-1}\right) = [\zeta(\theta_0 \mid \theta^{t-1}), ..., \zeta(\theta_N \mid \theta^{t-1})]$ in a $\mathcal{P}^t\left[\theta^{t-1}\right]$ consists of $N+1$ probability measures for each type $\theta_i$ given their past reports, $\theta^{t-1}$, and is thus closed. Consider $\mathcal{P} \equiv \bigcup_{t\in\mathbb{N}} \bigcup_{\theta^t \in \Theta^t} \mathcal{P}^t\left[\theta^{t-1}\right]$, which is a closed subset of $\mathcal{P}^{\infty}$. Since a closed subset of a compact space is compact (e.g., Dudley, 2002, Theorem 2.2.2), $\mathcal{P}$ is compact in the weak topology.

Finally, choosing $\phi \leq \beta$ in (B5) shows that the objective function is continuous in the weak topology. This establishes that including randomizations, we have a maximization problem over probability measures in which the objective function is continuous in the weak topology, and the constraint set is compact in the weak topology, and thus there exists a probability measure that reaches the maximum.

*Properties of* $\mathcal{U}(\{C_t, L_t\}_{t=0}^{\infty})$

We now establish the main properties of $\mathcal{U}(\{C_t, L_t\}_{t=0}^{\infty})$. The only additional restriction is that in all the proofs we assume that the solution to the maximization problem (3.10) is at a regular point. This needs to be imposed as an assumption, since it is not possible to check that the solution is indeed at a regular point. Nevertheless, this assumption is not a strong one, since if the solution is not at their regular point, a perturbation of the utility functions or the production function should ensure that the solution shifts to a regular point (i.e., solutions that are not at regular points in this context are "non-generic," though we do not present a precise mathematical statement of this property to economize on further notation and space).

**Lemma B2** $U(\{C_t, L_t\}_{t=0}^{\infty})$ *is continuous and concave on* $\Lambda^{\infty}$*, nondecreasing in* $C_s$ *and nonincreasing in* $L_s$ *for any* $s$ *and differentiable in* $\{C_t, L_t\}_{t=0}^{\infty}$.

*Proof.* The above argument established that in the problem of maximizing (3.10) subject to (3.11), (3.12), and (3.14) over probability measures, a maximum exists and $\mathcal{U}(\{C_t, L_t\}_{t=0}^{\infty})$ is therefore well defined.

To show concavity, consider $(C^0, L^0)$ and $(C^1, L^1)$ and corresponding $\zeta^0, \zeta^1$. We have

$$\begin{aligned}
&\int (u(c,l;\theta) - u(c,l;\hat{\theta}))\zeta^{\alpha}(d(c,l),\theta) \\
&= \alpha \int (u(c,l;\theta) - u(c,l;\hat{\theta}))\zeta^{0}(d(c,l),\theta) + (1-\alpha) \int (u(c,l;\theta) - u(c,l;\hat{\theta}))\zeta^{1}(d(c,l),\theta) \\
&\geq 0
\end{aligned}$$

In a similar way we can show that $\zeta^{\alpha}$ satisfies (3.11), (3.12), and (3.14), this convex combination is feasible and it gives the same utility as $\alpha\zeta^0 \cdot u(\theta) + (1-\alpha)\zeta^1 \cdot u(\theta)$.

Next, note that the constraint set expands if $C_s$ increases or $L_s$ decreases for any $s$, therefore $U$ must be weakly increasing in $C_s$ and weakly decreasing in $L_s$.

Finally, returning to the original topology, $\mathcal{U}(\{C_t, L_t\}_{t=0}^{\infty})$ is defined over a Banach space. Given the assumption that the solution to (3.10) is at a regular point, we can use Theorem B1 to conclude that $\mathcal{U}(\{C_t, L_t\}_{t=0}^{\infty})$ is differentiable in $\{C_t, L_t\}_{t=0}^{\infty}$, completing the proof. ■

**Lemma B3** $\Lambda^{\infty}$ *is compact and convex.*

*Proof.* (**Convexity**) Consider $\{C_t, L_t\}_{t=0}^{\infty}$ and $\{C_t', L_t'\}_{t=0}^{\infty} \in \Lambda^{\infty}$ and some $\zeta^0, \zeta^1$ feasible for $\{C_t, L_t\}_{t=0}^{\infty}$ and $\{C_t', L_t'\}_{t=0}^{\infty}$, respectively. Now for any $\alpha \in (0,1)$, $\zeta^{\alpha} \equiv \alpha\zeta^0 + (1-\alpha)\zeta^1$ is feasible for $(\alpha \{C_t, L_t\}_{t=0}^{\infty} + (1-\alpha)\{C_t', L_t'\}_{t=0}^{\infty})$, so that this set is non-empty. Moreover, since $\zeta^0, \zeta^1$ satisfy the incentive compatibility constraints, $\zeta^{\alpha}$ satisfies it as well. Similarly, $\zeta^{\alpha}$ satisfies the constraits on aggregate $\{C_t, L_t\}_{t=0}^{\infty}$.

(**Compactness**) For any sequence $\{C_t^n, L_t^n\}_{t=0}^{\infty} \in \Lambda^{\infty}, \{C_t^n, L_t^n\}_{t=0}^{\infty} \to \{C_t^{\infty}, L_t^{\infty}\}_{t=0}^{\infty}$, there exists a sequence $\{\zeta^n\}_{t=0}^{\infty}$ corresponding to $\{C_t^n, L_t^n\}_{t=0}^{\infty}$, such that $\zeta^n \to \zeta^{\infty}$, satisfying the incentive compatibility, aggregate constraints and feasibility, therefore $\{C_t^{\infty}, L_t^{\infty}\}_{t=0}^{\infty} \in \Lambda^{\infty}$ is closed. Boundedness follows from boundedness of $C$ and $L$. ■

## PROOF OF THEOREM ??

*Proof.* We showed above that, when randomizations are introduced, $\mathcal{U}(\{C_t, L_t\}_{t=0}^{\infty})$ is a well-defined functional and is continuous, concave, and differentiable. In this proof, we suppress randomization to simplify notation.

We write the problem of characterizing the best sustainable mechanism non-recursively following Marcet and Marimon (1998) as

$$\max_{\{C_t, L_t, K_t, x_t\}_{t=0}^{\infty}} \mathcal{L} = \mathcal{U}(\{C_t, L_t\}_{t=0}^{\infty}) + \sum_{t=0}^{\infty} \delta^t \left\{\mu_t v(x_t) - (\mu_t - \mu_{t-1}) v(F(K_t, L_t))\right\} \tag{B6}$$

subject to

$$C_t + x_t + K_{t+1} \leq F(K_t, L_t), \text{ and} \tag{B7}$$

$$\{C_t, L_t\}_{t=0}^{\infty} \in \Lambda^{\infty},$$

for all $t$, where $\mu_t = \mu_{t-1} + \psi_t$ with $\mu_{-1} = 0$ and $\delta^t \psi_t \geq 0$ is the Lagrange multiplier on the constraint (3.14). The differentiability of $\mathcal{U}(\{C_t, L_t\}_{t=0}^{\infty})$ implies that for $\{C_t, L_t\}_{t=0}^{\infty} \in \text{Int}\Lambda^{\infty}$, we have:

$$\mathcal{U}_{L_t} - \delta^t(\mu_t - \mu_{t-1})v'(F(K_t, L_t))F_{L_t} = -\mathcal{U}_{C_t} \cdot F_{L_t} \tag{B8}$$

$$\mathcal{U}_{C_t} = \left[\mathcal{U}_{C_{t+1}} - \delta^t(\mu_{t+1} - \mu_t)v'(F(K_{t+1}, L_{t+1}))\right] F_{K_{t+1}} \tag{B9}$$

Since $\mu_t \geq \mu_{t-1}$, there will be downward labor and intertemporal distortions whenever $\mu_t > \mu_{t-1}$ and $\mu_{t+1} > \mu_t$, respectively, i.e., whenever $\psi_t > 0$ and $\psi_{t+1} > 0$.

**Part 1:** Suppose to obtain a contradiction that $\mu_t = 0$ for all $t \geq 0$. Then, no consumption is allocated to the politician, $x_t = 0$ for all $t$. But in this case, if $L_t > 0$ for any $t$, then the politician can improve by expropriating the entire output at $t$. Thus we must have $L_t = 0$ for all $t$. Since, by hypothesis, $\{C_t, L_t\}_{t=0}^{\infty} \in \text{Int}\Lambda^{\infty}$ with $L_t > 0$ is feasible and the associated $\{C_t, L_t\}_{t=0}^{\infty} \in \text{Int}\Lambda^{\infty}$ necessarily gives higher ex ante utility to citizens than $L_t = C_t = 0$, the plan with $L_t = 0$ for all $t$ cannot be optimal. Therefore, the sustainability constraint of politician (3.14) must bind at some $t$ with $\psi_t > 0$. Then (B8) implies that there will be downward labor distortions at that $t$, and (B9) implies that there will be downward intertemporal distortions at $t-1$.

**Part 2:** We start by proving that $\varphi \equiv \inf\{\ \varrho \in [0,1] : \text{plim}_{t\to\infty} \varrho^{-t} U^*_{C_t} = 0\}$ is well-defined and strictly less than 1. To see this, recall that by hypothesis, a steady state exists, so that $\{C_t, L_t, K_{t+1}\}_{t=0}^{\infty} \to (C^*, L^*, K^*)$, thus $\{C_t\}_{t=0}^{\infty}$ is in the space $c$ of convergent infinite sequences (rather than simply in the space of all bounded infinite sequences, $\ell_\infty$). The dual of $c$ is $\ell_1$, that is, the space of sequences $\{y_t\}_{t=0}^{\infty}$ such that $\sum_{t=0}^{\infty} |y_t| < \infty$. Since $\mathcal{U}_{C_t}$ is equal to the Lagrange multiplier for the constraint (3.18), it lies in the dual space of $\{C_t\}_{t=0}^{\infty}$ (see, e.g., Luenberger, 1969, Chapter 9), thus in $\ell_1$, which implies that $\lim_{t\to\infty} \mathcal{U}_{C_t} = 0$, hence $\varphi < 1$.

Rearranging equations (B8) and (B9) and substituting for $\mathcal{U}_{C_t}$ and taking the limit as $t \to \infty$, we have

$$-\frac{\mathcal{U}^*_{L_t}}{\mathcal{U}^*_{C_t} F_{L_t}(K^*, L^*)} = 1 - \frac{(\mu_t - \mu_{t-1})v'(F(K^*, L^*)))}{\mu_t v'(x^*)} \tag{B10}$$

and

$$\frac{F_{K_{t+1}}(K^*, L^*)\mathcal{U}^*_{C_{t+1}}}{\mathcal{U}^*_{C_t}} = 1 + \frac{(\mu_{t+1} - \mu_t)v'(F(K^*, L^*))F_{K_{t+1}}(K^*, L^*)}{\mu_t v'(x^*)}, \tag{B11}$$

where all derivatives are evaluated at the limit $(C^*, L^*, K^*)$.

The first-order condition with respect to $x_t$ then implies:

$$\frac{\mathcal{U}_{C_t}}{\delta^t v'(x_t)} = \mu_t \leq \mu_{t+1} = \frac{\mathcal{U}_{C_{t+1}}}{\delta^{t+1} v'(x_{t+1})}. \tag{B12}$$

By construction, $\mu_t$ is an increasing sequence, so it must either converge to some value $\mu^*$ or go to infinity. Since as $t \to \infty$ an interior steady state $(C^*, L^*, K^*, x^*)$ exists by hypothesis and $\mathcal{U}^*_{C_t}$ is proportional to $\varphi^t$, (B12) can be written as

$$\frac{\varphi^t \mathcal{U}^*_{C^*}}{\delta^t v'(x^*)} = \mu_t \leq \mu_{t+1} = \frac{\varphi^{t+1} \mathcal{U}^*_{C^*}}{\delta^{t+1} v'(x^*)} \text{ as } t \to \infty. \tag{B13}$$

Since $\varphi = \delta$, we have that (B13) implies that as $t \to \infty$, $|\mu_{t+1} - \mu_t| \to 0$ and $\mu_t \to \mu^* \in (0, \infty]$ (where the fact that $\mu^* > 0$ follows from Part 1, since $\mu_{t+1} \geq \mu_t$ and $\mu_t > 0$ for some $t$). Therefore, $(\mu_t - \mu_{t-1})/\mu_t \to 0$, and distortions disappear asymptotically.

**Part 3:** Suppose that $\varphi > \delta$. In this case, (B12) implies that $\mathcal{U}^*_{C_t}$ is proportional to $\varphi^t$ as $t \to \infty$. This implies that $(\mu_t - \mu_{t-1})/\mu_t > 0$ as $t \to \infty$, so from (B8) and (B9), aggregate distortions cannot disappear, completing the proof. ■

## REFERENCES FOR APPENDIX B


**Bertsekas, Dimitri, Angelia Nedic and Asuman Ozdaglar (2003)** *Convex Analysis and Optimization*, Athena Scientific Boston.

**Dudley, Richard (2002)** *Real Analysis and Probability*, New York, Cambridge University Press.

**Luenberger, David G. (1969)** *Optimization by Vector Space Methods*, John Wiley & Sons New York.

**Marcet, Albert and Ramon Marimon (1998)** “Recursive Contracts” Mimeo. University of Pompeu Fabra

**Parthasarathy, K. R. (1967)** *Probity Measures on Metric Spaces*, Academic Press, New York.