Land Productivity and Statehood: The Surplus Theory Revisited

Nils-Petter Lagerlöf
Department of Economics
York University
www.nippelagerlof.com
Email: lagerlof@yorku.ca

April 2018

Abstract: The Surplus Theory attributes the rise of states and social stratification to an increase in land productivity following the Neolithic Revolution. The argument is that a non-producing elite could only be supported with ample enough food supply, i.e., a "surplus." It has also been argued that this theory is inconsistent with standard Malthusian theory, where land productivity has no effect on steady-state living standards, and thus generates no surplus in the first place. The current paper first illustrates this point in a simple Malthusian model, where a ruler chooses how much to extract from his subjects: in steady state the extraction rate is independent of land productivity. It then proposes a richer model, where the ruler can also undertake investments in productive and extractive capacities. In this extended model, gradual improvements in land productivity can trigger an endogenous transition from low to high extractive capacity, and a rise in output and population, essentially what the Surplus Theory posits.

1 Introduction

For most of its existence the human species has lived in small bands of hunters and gatherers. Organized, complex, and hierarchical social structures—what we often call *states*—are a relatively recent phenomenon. States emerged gradually from around 3500 BCE, starting in a few corners of the world, in particular Mesopotamia, China, the Nile and Indus Valleys, Mesoamerica, and the Andes (e.g., Service 1975, Ch. 1; Borcan et al. 2018). A few millennia earlier, these same regions were also the first to enter the Neolithic Revolution, i.e., develop agriculture.

Many have therefore hypothesized a causal link from land productivity to statehood. (See, e.g., Childe 1936; Diamond 1997; Hibbs and Olsson 2004; Putterman 2008, Section IV.) One proposed mechanism has been labelled the Surplus Theory. The idea is that agriculture caused, or allowed, the rise of states by raising output per unit of land, thus creating a "surplus" which could be stored, and then feed a ruling elite. By contrast, in human societies which rely on relatively low-yielding techniques to obtain food, no such elite population can be sustained, since everyone's labor is needed for procuring food.[1]

This Surplus Theory has also been critiqued. In particular, Mayshar et al. (2018) argue that the relevant innovation was not the rise in land productivity as such, but rather the properties of new crops, foremost cereals. These were easier to expropriate than foods obtained through gathering or horticulture, in particular tubers. Their case is mainly empirical, as they find scant evidence of earlier statehood in locations with higher agricultural yields overall, when controlling for the relative productivity of cereals and tubers. But they also make the theoretical point that the Surplus Theory is hard to reconcile with Malthusian theory. This is potentially important, given the broad consensus about the relevance of the Malthusian model for preindustrial development (see, e.g., Galor 2010, Ashraf and Galor 2011).

The current paper continues this ongoing struggle to make theoretical sense of the Surplus Theory. The starting point is a standard Malthusian growth model, similar to that considered by, e.g., Mayshar et al. (2018, Appendix B). Workers, or what is here called subjects, face a subsistence consumption constraint, and reproduction is proportional to income above subsistence, what may be called a "surplus." A ruler chooses how much to extract of

[1] For example, Hibbs and Olsson (2004, p. 3718) write that "[t]he superior agricultural mode of production made possible specialization of economic activity and the establishment of a non-food producing class devoted to the creation and codification of knowledge and the development of technology." Diamond (1997, p. 285) writes that "food production [i.e., agriculture] may be organized so as to generate stored food surpluses, which permit economic specialization and social stratification." In Mann (1986), an oft-cited overview of the literature on early state development, the index lists 26 pages referencing the term "surplus" in various contexts.

the subjects' incomes, taking into account how the future size of the subject population is affected.

We first show that the ruler's optimal extraction rate indeed increases following a one-time shock to land productivity, ostensibly consistent with the Surplus Theory. However, in a steady state with constant and non-growing levels of land productivity, the extraction rate is independent of productivity. The intuition relates to the standard Malthusian result that (pre-extraction) incomes are independent of land productivity, and the extraction rate here depends on productivity only through pre-extraction incomes. This illustrates how a simple Malthusian model can reject the Surplus Theory, confirming the result of Mayshar et al. (2018, Appendix B).

One novel finding is worth noting: an increase in the *growth rate* of land productivity (rather than its level) does indeed raise the extraction rate along a Malthusian balanced growth path. This is of potential interest, in as far as the Neolithic Revolution constituted a transition from slow to faster growth in land productivity, and not only a one-time rise in productivity levels.

At the more conceptual level, the setting described so far abstracts from a couple of mechanisms often discussed when linking land productivity to state development. First, it was assumed that the ruler consumes all he extracts, but early states were often instrumental in providing public goods. Typical examples include irrigation systems (cf. Wittfogel 1957, Nissen and Heine 2009), and external defense (cf. Dal Bó et al. 2016). The creation of new knowledge may also be a type of public goods investment. Second, by letting the ruler freely choose the extraction rate, we assumed away any state building process through which he accumulates power, or capacity, to extract resources (cf. Besley and Persson 2009, 2011).

This brings us to the main contribution of the current paper, namely a model where the ruler can undertake investments in productive and extractive capacities. This introduces richer dynamic links through which extraction in one period can impact that in the next. By investing in productive capacity, the ruler can raise future output and thus expand the future tax base. Investment in extractive capacity improves the efficiency of the extraction process itself, allowing the ruler to extract more resources at lower cost to the subjects' reproduction.

In this extended model, for certain levels of land productivity, there may be multiple steady state equilibria: one steady state has low extractive capacity and low levels of population and output; another has high extractive capacity and high population and output. Intuitively, more output provides stronger incentives for rulers to invest in extractive capacity, and greater extractive capacity sustains higher output by enabling investments in productive capacity.

A temporary shock to output, or to extractive capacity, can push the economy from one steady state to another, even while holding the land productivity parameter constant. Moreover, a permanent increase in land productivity can alter the configuration of the dynamical system, such that the low-extractive steady state goes away, and only the high-extractive one remains. As a corollary, slow and gradual improvements in land productivity can at some point trigger an endogenous transition from low to high extractive capacity, and an associated rise in output and population, thus capturing the main point of the Surplus Theory.

This paper seeks to contribute to a broad literature on early state development. One strand of that literature focuses on the consequences of early statehood for modern development. For example, Borcan et al. (2018) document that countries with very early and very late statehood tend to do worse than those with states of intermediate age. Other studies using earlier installments of those state antiquity data (e.g., Bockstette et al. 2002, Chanda and Putterman 2007, Chanda et al. 2014) find a mostly positive relationship. There are also some interesting correlations between early statehood and other modern outcome variables: Hariri (2012) documents that countries with older states are currently less democratic; Depetris-Chauvin (2016) finds links between early statehood and modern conflict in Africa. Theories linking the timing of statehood to democracy and modern development include Lagerlöf (2016).

Theories on why states emerged when and where they did include Mayshar et al. (2017, 2018), who emphasize the appropriability of resources (see discussion above). While more modest in scope, the current paper incorporates some mechanisms related to their so-called Appropriability Theory. In particular, a shock to extractive capacity—which in a sense captures the appropriability of resources—can push a society from a low- to high-extractive steady state, holding constant land productivity. One important difference is that we here use a Malthusian setting throughout, while the baseline framework used by Mayshar et al. (2017, 2018) mostly treat population as exogenous.[2]

Other theories of the origin of statehood include that of Dal Bó et al. (2016), who focus on the state's role as protector. One result of their model is that states are more likely to emerge when the natural environment provides more protection. This in a sense captures the interaction between what we may call productive and defensive capacities, while we here focus on productive and extractive capacities. Conceptually, extractive capacity could represent the powers of a domestic and forward-looking ruler with some incentives to invest in productive capacity. By contrast, defensive capacity would capture the ability to protect against extraction by external and less benevolent actors. Another difference compared to Dal Bó et al. (2016) is that we here apply a Malthusian framework, while they abstract from population dynamics.[3]

[2] One noteable exception is the model in Mayshar et al. (2018, Appendix B).

[3] There are also theories of state emergence emphasizing ecological conditions that promote trade and

Finally, this paper leans on a theoretical literature, in particular by Besley and Persson (2009, 2011), on fiscal and legal state capacities and their complementarities; what we here call extractive capacity corresponds closest to fiscal capacity in their jargon. Somewhat different from the baseline Besley-Persson framework, we model extractive capacity as a ruler's power over tax collectors. Another difference is that we use a dynamic setting, where investments in public goods impact the future tax base.[4] This puts our focus more squarely on the interplay between productive and extractive capacities, similar to, e.g., Acemoglu (2005). Yet again, none of these papers uses a Malthusian framework.

The rest of this paper is organized as follows. Section 2 illustrates the fallacy of the Surplus Theory in a simple Malthusian model, where agents face a subsistence consumption constraint, and where investments in extractive or productive capacities are absent. Section 3 extends the setting in Section 2 to allow for such investments, leading up to a couple of results that seem more in line with the Surplus Theory. Section 4 ends with a concluding discussion.

2 Model A: subsistence consumption

This section presents a basic model to illustrate one simple interpretation of the Surplus Theory, and its potential shortcomings. The central model components are a subsistence consumption constraint on part of the subject population, and a ruler who chooses how much to extract from his subjects.

Consider thus a world where a class of subjects live in overlapping generations for two periods: as passive children and active adults. Part of their income is expropriated by a ruler. We refer to the ruler in the singular, and using the male pronoun, although this could be a group of agents. In the adult phase of life, a subject works, pays taxes, and produces offspring, making the size of the subject population evolve endogenously over time, as a function of the ruler's extraction rate. The ruler, by contrast, does not work, which can be interpreted as him having one offspring, replacing the (single) parent in the next period.

2.1 Production and productivity growth

Output in period t, denoted Y_t, is produced using the technology

$$Y_t = (MB_t)^{\alpha} L_t^{1-\alpha}, \tag{1}$$

specialization (see, e.g., Fenske 2014, Depetris-Chauvin and Özak 2016). However, this literature is mostly empirical, and may be more tangetially related to the current exercise.

[4] Besley et al. (2013) provide an example of a dynamic model of investment in state capacity.

where α is the land share of output, L_t is the size of the (working) population of subjects (each suppling one unit of labor), and M is the size of land. For the rest of the presentation we normalize land to unity ($M = 1$). The factor B_t denotes land productivity, which in this setting grows at the exogenous rate $g \geq 0$, i.e.,

$$B_{t+1} = (1 + g)B_t. \tag{2}$$

This formulation nests the special case when $g = 0$, implying that B_t is constant at its initial level, B_0. This allows us to do comparative statics with respect to B_0.

2.2 Extraction and population dynamics

The ruler has the power to tax (i.e., extract resources from) the subjects, which comes at the cost of lowering their reproduction and thus the ruler's future tax revenues. To capture this, let each subject earn income $y_t = Y_t/L_t = (B_t/L_t)^\alpha$, which is taxed at rate $\tau_t \in [0, 1]$. Each subject thus earns $(1 - \tau_t)y_t$ after taxes. Subjects care about consumption, c_t^S, and fertility, n_t. Utility is given by

$$U_t^S = (1 - \widetilde{\gamma}) \ln \left(c_t^S - \widetilde{c}\right) + \widetilde{\gamma} \ln n_t, \tag{3}$$

where $\widetilde{\gamma} \in (0, 1)$, and where $\widetilde{c} \geq 0$ denotes subsistence consumption, below which utility is not defined.

Each subject takes her income as given and maximizes (3) subject to the budget constraint

$$c_t = (1 - \tau_t)y_t - qn_t, \tag{4}$$

where $q > 0$ is the (goods) cost per child. This gives optimal fertility as

$$n_t = \gamma \left[(1 - \tau_t)y_t - \widetilde{c}\right]. \tag{5}$$

where $\gamma \equiv \widetilde{\gamma}/q$. Fertility thus approaches zero as post-extraction income per subject approaches subsistence.

Since each subject is replaced by n_t offspring, the subject population in the next period equals $L_{t+1} = n_t L_t$. Applying (5) and $y_t = Y_t/L_t$ gives

$$L_{t+1} = \gamma \left[(1 - \tau_t)Y_t - \widetilde{c}L_t\right] = \gamma \left[(1 - \tau_t)y_t - \widetilde{c}\right] L_t. \tag{6}$$

That is, the population of subjects in the next period, L_{t+1}, decreases with the ruler's current rate of extraction, τ_t.

2.3 The ruler's optimal extraction rate

The ruler's income is $\tau_t Y_t$, of which he consumes all. His preferences are defined over his consumption, $\tau_t Y_t$, and the total tax base in the next period, Y_{t+1}, and are described by the utility function

$$U_t^R = (1-\beta)\ln(\tau_t Y_t) + \beta \ln(Y_{t+1}), \tag{7}$$

where $\beta \in (0,1)$.

The ruler sets his only choice variable, τ_t, to maximize utility in (7), subject to (1) forwarded one period [i.e., $Y_{t+1} = B_{t+1}^{\alpha} L_{t+1}^{1-\alpha}$], and the dynamics of the subject population, as described by (6). The first-order condition states that

$$(1-\beta)\left[\frac{1}{\tau_t}\right] = \beta(1-\alpha)\left[\frac{y_t}{(1-\tau_t)y_t - \widetilde{c}}\right], \tag{8}$$

i.e., the marginal value of extraction equals the marginal loss of future tax revenue, working through the effect of taxation on subjects' fertility, and thus on the next period's workforce. Solving (8) for τ_t gives

$$\tau_t = \left(\frac{1-\beta}{1-\alpha\beta}\right)\left(1 - \frac{\widetilde{c}}{y_t}\right). \tag{9}$$

This arguably captures the essence of the Surplus Theory: the rate at which the ruler taxes his subjects, τ_t, is increasing with pre-extraction per-capita incomes, y_t. Indeed, a positive extraction rate requires a "surplus," i.e., a gap between income per subject and subsistence ($y_t > \widetilde{c}$). Note also that the positive relationship between τ_t and y_t in (9) hinges on the presence of a subsistence consumption constraint, $\widetilde{c} > 0$. If $\widetilde{c} = 0$, then τ_t is constant at $(1-\beta)/(1-\alpha\beta)$ and independent of y_t.

However, (9) describes an out-of-steady-state relationship between two endogenous variables. Next we study the dynamics of y_t, and what determines its steady state level, and thus also the steady-state extraction rate.

2.4 Dynamics and steady state

Using (2), (6), (9), and $y_{t+1} = (B_{t+1}/L_{t+1})^{\alpha}$, we can derive an expression for the gross growth rate in (pre-extraction) per-capita income (y_{t+1}/y_t) in terms of its initial level (y_t):

$$\frac{y_{t+1}}{y_t} = \left[\frac{(1-\alpha\beta)(1+g)}{\gamma\beta(1-\alpha)}\right]^{\alpha}\left(\frac{1}{y_t - \widetilde{c}}\right)^{\alpha} \equiv \Gamma(y_t). \tag{10}$$

This is illustrated in the bottom panel of Figure 1. It is easy to verify that $\Gamma'(y_t) < 0$, $\lim_{y_t \to \widetilde{c}} \Gamma(y_t) = \infty$, and $\lim_{y_t \to \infty} \Gamma(y_t) = 0$. This implies a unique steady state level of y_t, here denoted y^*, defined by $\Gamma(y^*) = 1$. (As always when imposing a subsistence consumption constraint in a Malthusian setting, it can also be shown that the dynamics around the steady state may be oscillatory. However, this is not of major interest here.) Setting $\Gamma(y^*) = 1$ in (10) gives

$$y^* = \left[\frac{(1-\alpha\beta)(1+g)}{\gamma\beta(1-\alpha)}\right] + \widetilde{c}. \tag{11}$$

Evaluating (9) at $y_t = y^*$, the steady-state level of the extraction rate, denoted τ^*, is given by

$$\tau^* = \frac{(1-\beta)(1+g)}{\widetilde{c}\beta\gamma(1-\alpha) + (1-\alpha\beta)(1+g)}, \tag{12}$$

which is illustrated in the top panel of Figure 1.

Recall that the model nests constant land productivity as a special case. That is, setting $g = 0$ in (2) implies $B_t = B_0$ for all $t \geq 0$. We can now state the following.

Proposition 1 *Consider Model A with* $\widetilde{c} > 0$.

(a) *Suppose productivity is non-growing ($g = 0$). Then the following holds:*

(i) A one-time increase to productivity (a jump in B_0 in period t), holding population (L_t) constant, raises per-worker output (y_t) and the extraction rate (τ_t) in the same period.

(ii) Steady-state per-worker output (y^*) and the steady-state extraction rate (τ^*) are independent of the level of productivity (B_0).

(b) *If productivity is growing ($g > 0$), then both steady-state per-worker output (y^*) and the steady-state extraction rate (τ^*) are increasing in the productivity growth rate (g).*

Proof: Part (a): The result in (i) follows from setting $y_t = (B_0/L_t)^\alpha$, which is increasing in B_0 when holding L_t constant; then (9) shows that τ_t is increasing in B_0. The result in (ii) follows directly from setting $g = 0$ in (11) and (12), which shows that y^* and τ^* are independent of B_0. Part (b): The result follows directly from (11) and (12), which show that y^* and τ^* are increasing in g. ∎

Part (a) of Proposition 1 illustrates the fallacy of the Surplus Theory in its simplest interpretation: the positive out-of-steady state relationship between land productivity and the ruler's optimal extraction rate goes away when imposing steady state. Part (b) suggests that, if we instead interpret the Neolithic Revolution as a rise in the sustained growth rate of land productivity, then the model's predictions are more consistent with the Surplus Theory. This is a useful insight in itself, if the only aim is a simple way to make sense of the Surplus Theory in a Malthusian framework.

These results are also quite robust. For example, it was assumed that the ruler extracts output (i.e., food), while some rather describe the surplus in terms of labor (see, e.g., Allen 1997, pp. 139-141, in the context of Egypt). However, nothing changes qualitatively if the ruler instead extracts labor services from his subjects.[5]

[5] To see this, let output equal $Y_t = B_t^\alpha([1-u_t]L_t)^{1-\alpha}$, where u_t is the fraction subjects serving the ruler.

At the same time, there are other problems with this framework as interpretation of the Surplus Theory. First, the ruler consumes all resources he extracts, so the model lacks any feedback from current extraction to future production. This is a problematic, since many authors linking state building to agriculture, and the concept of a surplus, emphasize the use of extracted resources for building public goods, such as irrigation systems (Wittfogel 1957), and the development of writing, science, and mathematics (e.g., Childe 1936, Ch. VIII; Hibbs and Olsson 2004, p. 3178).

Second, the ruler's ability, or power, to extract resources is not modelled, but statehood is rather interpreted as his chosen rate of extraction. In other words, the model lacks anything that can be interpreted as a *state building* process, i.e., an endogenous accumulation of the ruler's powers to extract resources. One early concrete example of such state building efforts could be the training of record-keeping scribes and bureaucrats. Schools were founded for that purpose in Mesopotamia around 2000 BCE, during the Third Dynasty of Ur (Nissen and Heine 2009, Ch. 5).

The model explored in the next section endogenizes investments in both extractive and productive capacities.

3 Model B: extractive and productive capacities

Consider now a model similar to Model A in Section 2, but with three classes: subjects, a ruler (as before referred to in the singular), and now also tax collectors. The tax collectors collect taxes from subjects on behalf of the ruler, and are able to keep a fraction $1 - z_t$ of the tax revenue, where $z_t \in [0, 1]$ is referred to as the ruler's *extractive capacity*. In short, z_t measures how much of the collected taxes end up in the ruler's coffers. As in Model A, the ruler still sets the rate of extraction, τ_t, but now keeps only a fraction z_t.[6]

Both the ruler and the tax collectors are of fixed size, while the subject population evolves endogenously in a Malthusian fashion, similar to Model A.

Moreover, let the ruler's utility be defined over current services received, $u_t L_t$, and the future tax base, here L_{t+1}, with the utility function $U_t^R = (1-\beta)\ln(u_t L_t) + \beta \ln(L_{t+1})$. We can then no longer solve analytically for the ruler's optimal choice of u_t, but it can be seen to be increasing in y_t, similar to how the food extraction rate τ_t does in (9). Moreover, the steady-state level of u_t, call it u^*, turns out to be independent of the (non-growing) level of land productivity B_0.

[6] Here this could capture the idea that the ruler can dominate either the subjects by allying with the tax collectors (who always prefer higher tax rates, since they keep some of the revenue), or the tax collectors by allying with the subjects (who prefer lower taxes).

3.1 Production and productivity

Output in period t, denoted Y_t, is produced using a production function similar to that in (1), but now with two sources of land productivity:

$$Y_t = (A_t B_t)^{\alpha} L_t^{1-\alpha}. \tag{13}$$

As before, α is the land share of output, and L_t is the size of the subject population. (The size of land is again normalized to one, $M = 1$.) B_t and A_t are the two land productivity factors. As in Model A, B_t is taken as given by the ruler. It could here capture both time-invariant factors, such as the geographical environment (e.g., the number of domesticable plants and animal species), and any productivity growth occurring absent directed investments undertaken by the ruler (e.g., random discoveries leading to improvements in agricultural productivity, or new crops arriving from other regions). By contrast, A_t depends on productivity-enhancing investments undertaken by the ruler, representing public goods such as irrigation systems, or knowledge. We shall refer to A_t as *productive capacity*.

As we shall see, one feature of this model is that slow and gradual improvements in B_t at some point can trigger a transition that can be interpreted as a state building process.

3.2 Population dynamics

Different from Model A, we here abstract from subsistence consumption, setting $\widetilde{c} = 0$. The reason for this is twofold. First, in Model A we noted that $\widetilde{c} > 0$ was a necessary condition for a positive relationship between τ_t and y_t to hold even outside steady state. Here subsistence consumption is not necessary to generate such a relationship, and would rather obscure the mechanisms which actually give rise to it. Second, as mentioned already, introducing subsistence consumption in a Malthusian model can give rise to oscillatory dynamics. While this can be interesting in its own right, it has nothing to do with extraction per se, thus being something of a red herring.

Therefore, we here let a subject's utility be given by (3), but now with $\widetilde{c} = 0$. Each subject faces a budget constraint like that in (4). Utility maximization thus leads to fertility as in (5), and population dynamics as in (6), but with $\widetilde{c} = 0$, i.e.,

$$L_{t+1} = \gamma(1 - \tau_t) y_t L_t = \gamma(1 - \tau_t) Y_t, \tag{14}$$

where (as in Model A) γ depends on the fertility weight in the utility function and the cost of children, and where $y_t = Y_t / L_t = (A_t B_t / L_t)^{\alpha}$; see (13).

3.3 The ruler's decision problem

Recall that the ruler's share of total tax revenues is z_t, the reminder being purloined by the tax collectors. The collective income of the tax collector class thus equals $\tau_t(1 - z_t)Y_t$, all of which is consumed and plays no role for the rest of the analysis, while the ruler's income is $\tau_t z_t Y_t$. We shall refer to $z_t Y_t$ as the ruler's *effective tax base.*

Let the ruler's investment in next period's extractive capacity be denoted $x_t \geq 0$, which builds extractive capacity in the next period, z_{t+1}, at a rate $\phi > 0$. We let extractive capacity be bounded from above and below at levels $\overline{z}$ and $\underline{z}$, respectively, such that $0 < \underline{z} < \overline{z} \leq 1$. More precisely,

$$z_{t+1} = \min\{\overline{z}, \underline{z} + \phi x_t\} = \begin{cases} \overline{z} & \text{if} \quad x_t \geq \frac{\overline{z}-\underline{z}}{\phi}, \\ \underline{z} + \phi x_t & \text{if} \quad x_t \in \left(0, \frac{\overline{z}-\underline{z}}{\phi}\right), \\ \underline{z} & \text{if} \quad x_t = 0. \end{cases} \tag{15}$$

Assuming some minimum level of extractive capacity ($\underline{z} > 0$) is important, because it ensures that the ruler always has some revenue with which to begin accumulating further extractive and productive capacity. This may capture that many egalitarian and pre-agrarian societies tend to have a "big man," a sort of embryonic chief or ruler (Read 1959; Service 1975; Ch. 4). Assuming $\overline{z} < (\leq)1$ is unimportant, and only serves to allow tax collectors some strictly positive (or non-negative) income.

Consider next investments in productive capacity. We let the cost of A_{t+1} in terms of period-t consumption be ηA_{t+1}^{σ}, where $\eta > 0$ and $\sigma > 1$. Assuming $\sigma > 1$ ensures that output and population converge to constant non-growing levels, absent growth in B_t.

The budget constraint can now be written

$$c_t^R = \tau_t z_t Y_t - \eta A_{t+1}^{\sigma} - x_t, \tag{16}$$

where c_t^R is the ruler's consumption. Analogously to Model A, the ruler's preferences are defined over c_t^R and the total effective tax base in the next period, $z_{t+1}Y_{t+1}$, with utility function

$$U_t^R = (1 - \beta) \ln\left(c_t^R\right) + \beta \ln(z_{t+1}Y_{t+1}), \tag{17}$$

where $\beta \in (0, 1)$. The ruler sets A_{t+1}, z_{t+1}, and τ_t to maximize (17), subject to (14), (15), (16), and (13) forwarded one period. Section A of the appendix derives the optimal choice of these as functions of the effective tax base, $z_t Y_t$, and exogenous parameters.

First let $\underline{X}$ and $\overline{X}$ denote the thresholds for $z_t Y_t$, above and below which the two constraints on z_{t+1} bind. These thresholds are given by

$$\overline{X} = \frac{1}{\phi}\left[\overline{z}\left(\frac{\beta\sigma(1-\alpha) + \sigma + \alpha\beta}{\beta\sigma}\right) - \underline{z}\right], \tag{18}$$

and

$$\underline{X} = \frac{1}{\phi}\left(\frac{\sigma(1-\alpha\beta)+\alpha\beta}{\beta\sigma}\right)\underline{z}. \tag{19}$$

It is straightforward to show that $0 < \underline{X} < \overline{X}$ follows from $0 < \underline{z} < \overline{z}$. The ruler's optimal extraction rate can now be written:

$$\tau_t = \begin{cases} 1 - \left[\frac{\beta\sigma(1-\alpha)}{\sigma(1-\alpha\beta)+\alpha\beta}\right]\left[1 - \left(\frac{\overline{z}-\underline{z}}{\phi}\right)\frac{1}{z_t Y_t}\right] & \text{if} \quad z_t Y_t \geq \overline{X}, \\ 1 - \left(\frac{\beta\sigma(1-\alpha)}{\beta\sigma(1-\alpha)+\sigma+\alpha\beta}\right)\left(1 + \frac{\underline{z}}{\phi z_t Y_t}\right) & \text{if} \quad z_t Y_t \in \left[\underline{X}, \overline{X}\right], \\ \frac{\sigma(1-\beta)+\alpha\beta}{\sigma(1-\alpha\beta)+\alpha\beta} & \text{if} \quad z_t Y_t \leq \underline{X}. \end{cases} \tag{20}$$

It is easy to see from (20) that τ_t is increasing in $z_t Y_t$ for $z_t Y_t \in \left[\underline{X}, \overline{X}\right]$. In this interval, rulers are building extractive capacity, and they respond to a larger effective tax base ($z_t Y_t$) by extracting more resources, in order to fund investment in future extractive capacity. We can call this a state building phase.

It can also be seen that τ_t decreases with $z_t Y_t$ for $z_t Y_t \geq \overline{X}$. When the ruler has reached maximum extractive capacity, he invests more in the future labor force by extracting fewer resources from the subjects. As $z_t Y_t$ approaches infinity, the extraction rate approaches the same level as when $z_t Y_t \leq \underline{X}$, i.e., $[\sigma(1-\beta)+\alpha\beta]/[\sigma(1-\alpha\beta)+\alpha\beta]$. Intuitively, the cost of maintaining extractive capacity becomes negligible relative to income as the effective tax base grows.

3.4 Dynamics

Since the optimal extraction rate in (20) depends on the effective tax base, $z_t Y_t$, the dynamics of the economy are most easily described in terms of the two state variables Y_t and z_t.

3.4.1 Dynamics of z_t

As shown in Section A of the appendix, the optimal choice of z_{t+1} can be written

$$z_{t+1} = \Phi(Y_t, z_t) \equiv \begin{cases} \overline{z} & \text{if} \quad z_t Y_t \geq \overline{X}, \\ \left(\frac{\beta\sigma}{\beta\sigma(1-\alpha)+\sigma+\alpha\beta}\right)\left[\phi z_t Y_t + \underline{z}\right] & \text{if} \quad z_t Y_t \in \left[\underline{X}, \overline{X}\right], \\ \underline{z} & \text{if} \quad z_t Y_t \leq \underline{X}. \end{cases} \tag{21}$$

That is, $z_{t+1} \geq \underline{z}$ binds when $z_t Y_t < \underline{X}$, and $z_{t+1} \leq \overline{z}$ binds when $z_t Y_t > \overline{X}$. When these constraints are non-binding—i.e., when $z_t Y_t \in \left[\underline{X}, \overline{X}\right]$—the solution is found in (A14) in the appendix. Over this interval, rulers with higher effective tax base invest more in future extractive capacity. (It is also easy to verify that the respective corner solutions coincide with the interior solution when $z_t Y_t = \underline{X}$, and $z_t Y_t = \overline{X}$.)

3.4.2 Dynamics of Y_t

Given the ruler's optimal choices for A_{t+1} and τ_t, we can derive an expression for Y_{t+1} in terms of z_t, Y_t, and B_{t+1}. Section B of the appendix shows that

$$Y_{t+1} = \Psi(Y_t, z_t, B_{t+1}) \equiv \begin{cases} \kappa D B_{t+1}^{\alpha} z_t^{\alpha-1} \left[\phi z_t Y_t + \underline{z} - \overline{z}\right]^{\rho} & \text{if} \quad z_t Y_t \geq \overline{X}, \\ D B_{t+1}^{\alpha} z_t^{\alpha-1} \left[\phi z_t Y_t + \underline{z}\right]^{\rho} & \text{if} \quad z_t Y_t \in \left[\underline{X}, \overline{X}\right], \\ \kappa D B_{t+1}^{\alpha} \left(\phi z_t Y_t\right)^{\rho} & \text{if} \quad z_t Y_t \leq \underline{X}, \end{cases} \tag{22}$$

where $\rho = (\alpha/\sigma) + 1 - \alpha < 1$, and where $D > 0$ and $\kappa > 1$ depend only on the exogenous and time-invariant variables α, β, γ, ϕ, σ, and η [see (A20) and (A27) in the appendix], and play no role for the dynamics.

Note that Y_{t+1} depends on B_{t+1}, i.e., the land productivity factor that is independent of the ruler's investments. In this section, we have not yet made any explicit assumptions about the dynamics of B_t, but an interesting feature of the model is that some very slow growth in B_t can trigger a growth spurt by inducing investments in z_t and A_t. To simplify the presentation, below we first let B_t be constant, and study how the economy responds to changes in its level. From there, it is easier to understand the full dynamics when we let B_t evolve endogenously.

3.4.3 Constant B_t

Without growth in B_t [i.e., setting $g = 0$ in (2)] it follows that $B_t = B_0$ for all $t \geq 0$. Now (21) and (22) define a two-dimensional dynamical system for z_t and Y_t:

$$\begin{aligned} z_{t+1} &= \Phi(Y_t, z_t), \\ Y_{t+1} &= \Psi(Y_t, z_t, B_0). \end{aligned} \tag{23}$$

The dynamical system in (23) is illustrated in the phase diagram in Figure 2. It shows the loci along which z_t and Y_t are constant (derived in Section C of the appendix), and the regions where the constraints on extractive capacity investments bind: $z_{t+1} \geq \underline{z}$ binds when $z_t Y_t < \underline{X}$, and $z_{t+1} \leq \overline{z}$ binds when $z_t Y_t > \overline{X}$.

Generally, the configuration depends on exogenous variables (in particular B_0, as discussed below). Figure 2 illustrates a case where there are two locally stable steady-state equilibria, and one unstable. (Existence conditions are given in Proposition 2 below.) One stable steady-state equilibrium can be labelled a *low-extractive* steady state. Here the ruler undertakes no investments in extractive capacity, so $z_t = \underline{z}$, and output can be written

$$\underline{Y} = \left[\kappa D B_0^{\alpha} \left(\phi \underline{z}\right)^{\rho}\right]^{\frac{1}{1-\rho}}, \tag{24}$$

which is illustrated in Figure 2, and derived by setting $Y_{t+1} = Y_t = \underline{Y}$ and $z_t = \underline{z}$ in the bottom row of (22). The associated extraction rate, which we can denote $\underline{\tau}$, is given by the bottom row of (20), i.e., $\underline{\tau} = [\sigma(1-\beta) + \alpha\beta]/[\sigma(1-\alpha\beta) + \alpha\beta]$. Population is given by (14) as $\underline{L} = \gamma(1-\underline{\tau})\underline{Y}$.

The other stable steady state, at which $z_t = \overline{z}$, can be labelled the *high-extractive* steady state. Here output equals $\overline{Y} > \underline{Y}$, where $\overline{Y}$ is defined from $\overline{Y} = \kappa D B_0^{\alpha} \overline{z}^{\alpha-1} \left[\phi \overline{z}\overline{Y} + \underline{z} - \overline{z}\right]^{\rho}$; cf. the top row of (22). The extraction rate in this steady state, $\overline{\tau}$, is given by the top row of (20), setting $z_t Y_t = \overline{z}\overline{Y}$. From (14), population can be written $\overline{L} = \gamma(1-\overline{\tau})\overline{Y}$.

Note that the steady-state levels of population, output, extractive capacity, and the rate of extraction are all endogenous and jointly determined. The following proposition sums up the main properties of the two steady states, and provides conditions for their respective existence.

Proposition 2 *Consider Model B, with $B_t = B_0$ for all $t \geq 0$. There exist $\widehat{B} > 0$ and $\widehat{\widehat{B}} > 0$, such that:*

(a) *If $B_0 < \widehat{B}$, then there exists a low-extractive steady state, $(\underline{z}, \underline{Y})$, such that $\underline{z}\underline{Y} < \underline{X}$.*

(b) *If $B_0 > \widehat{\widehat{B}}$, then there exists a high-extractive steady state, $(\overline{z}, \overline{Y})$, such that $\overline{z}\overline{Y} > \overline{X}$.*

(c) *For $\underline{z}$ small enough, it holds that $\widehat{\widehat{B}} < \widehat{B}$. That is, the low- and the high-extractive steady states can coexist for $B_0 \in (\widehat{\widehat{B}}, \widehat{B})$.*

(d) *If both steady states exist, then the following holds:*

(i) The low-extractive steady state has lower output than the high-extractive steady state, i.e., $\underline{Y} < \overline{Y}$;

(ii) The low-extractive steady state has lower population than the high-extractive steady state, i.e., $\underline{L} < \overline{L}$;

(iii) The low-extractive steady state has a lower extraction rate than the high-extractive steady state, i.e., $\underline{\tau} < \overline{\tau}$.

The proof is in Section D of the appendix.

The intuition behind this multiplicity of steady states has to do with the complementarity between extractive and productive capacities. When output is high the ruler has strong incentives to invest in extractive capacity to increase his share of the pie. To help finance these investments he chooses a higher rate of extraction; cf. (20). Greater extractive capacity in turn sustains tax revenues which the ruler can use for investments in productive capacity.

Note that the results under (d) in Proposition 2 are far from trivial. For example, a higher rate of extraction would seem to imply a smaller population keeping output constant; cf. (14). The fact that the high-extractive steady state still has larger population reflects the fact that it has higher output, due to higher investments in productive capacities, in turn sustained by the ruler's larger tax revenues.

Shocks to z_t or Y_t Given a configuration with multiple steady states, such as that in Figure 2, the economy converges over time to one of the stable ones. Which one it converges to depends on its initial position relative to the saddle-path trajectory leading to the unstable steady state.

This means that an economy can transit from the low- to the high-extractive steady state in the wake of a one-time shock to either extractive capacity (z_t), or output (Y_t), or a combination of the two. Intuitively, the one-time shock raises the ruler's income in period t, who then invests more in productive and/or extractive capacity, possibly putting the economy on a trajectory leading to the high-extractive steady state. For this to happen, the shock must push (z_t, Y_t) above the threshold saddle path.

A transition due to a shock to output would be consistent with the Surplus Theory, and could perhaps be interpreted as the result of temporary climatic variations, and/or a phase of good harvests. A transition due to a shock to extractive capacity relates conceptually to the theory of Mayshar et al. (2017, 2018) about how states arise when and where it is easy to appropriate resources. In our framework, this could be interpreted as the introduction of new crops, such as cereals, which are easier to store and/or confiscate than previously used crops.

Exogenous changes to B_0 Another way to think about the Surplus Theory in this framework is in terms of a gradual increases in B_0. As shown in Section C.3 of the appendix, this shifts up the $(Y_{t+1} = Y_t)$-locus, thus raising output in the low-extractive steady state; note from (24) that $\underline{Y}$ is increasing in B_0. It also expands the basin of attraction for the high-extractive steady state. At some point the low-extractive steady state ceases to exist. The economy then goes through a rapid spurt in extractive capacity and output, stabilizing at $\overline{z}$ and $\overline{Y}$, respectively, from which point on $\overline{Y}$ keeps expanding with B_0.

3.4.4 Endogenously evolving B_t

To take this one step further, we can postulate that B_t grows over time as a function of output, Y_t. Here we let

$$B_{t+1} = [1 + \min\{\overline{g}, \theta Y_t\}]B_t \equiv \Lambda(Y_t, B_t), \tag{25}$$

for some $\overline{g} > 0$, $\theta > 0$, and $\theta \approx 0$. In words, growth in B_t increases with Y_t, but only up to some maximum level $\overline{g}$. Thus, when output levels are low, $Y_t < \overline{g}/\theta$, the growth rate of B_t is very small, but positive. The full dynamical system is now three-dimensional, and defined by (21), (22), and (25):

$$\begin{aligned} z_{t+1} &= \Phi(Y_t, z_t), \\ Y_{t+1} &= \Psi(Y_t, z_t, \Lambda(Y_t, B_t)), \\ B_{t+1} &= \Lambda(Y_t, B_t). \end{aligned} \tag{26}$$

Figure 3 shows the time paths for select variables from a simulation of the dynamical system in (26) for a somewhat arbitrary set of parameter values. Initial conditions are set so that $z_0 = \underline{z}$, and $z_0 Y_0 = \underline{z} Y_0 < \underline{X}$, meaning the ruler initially invests nothing in extractive capacity, which is thus constrained to its minimum level, $\underline{z}$.

Because B_t is growing, so is Y_t. When Y_t exceeds $\underline{X}/\underline{z}$, the ruler starts to invest in extractive capacity, inducing higher extraction (τ_t), as well as growth in z_t, as shown in Panel A. This is the state building phase.

The ruler also increases his investment in productive capacity (A_{t+1}), which generates a growth spurt in output, and expands the ruler's income further; see Panels B and D. (Annual rates are calculated assuming that each model period is 25 years.)

As seen in Panel A, extractive capacity at some point reaches its maximum, $\overline{z}$. From then on, the extraction rate declines, and population growth spurts (Panel D). As seen from the perspective of the ruler, who cares about next period's effective tax base ($z_{t+1}Y_{t+1}$), this reflects a substitution from investment in extractive capacity to investment in the future labor force; the latter entails extracting less resources from the subjects.

Panel C of Figure 3 also shows how the relative income share of each class evolves over time. During the state building phase, subjects are taxed more, thus seeing a drop in their share. While the ruler is still relatively weak, this rise in extraction benefits tax collectors. As the ruler accumulates more extractive capacity, the tax collectors' share falls.

4 Concluding remarks

The Surplus Theory may be one of the most common explanations of the origin of statehood, and social stratification more generally. The idea is that a non-producing elite could only be supported with a "surplus" supply of food. This surplus, goes the argument, arrived when land productivity rose in the wake of the Neolithic Revolution, i.e., when humans transitioned from food procurement through hunting and gathering to using agriculture.

To make sense of the Surplus Theory within a simple Malthusian growth model is not easy. This paper starts off with a model where a ruler can choose how much to extract from his working subjects. The subjects face a subsistence consumption constraint, and the ruler extracts resources from his subjects taking into account how the size of the future labor force is impacted. Outside of steady state, it is straightforward to show that optimal extraction rate is increasing in land productivity. That is, a one-time shock to productivity raises the rate of extraction. However, in steady state the extraction rate is independent of productivity, contradicting the Surplus Theory.

To salvage the Surplus Theory in a Malthusian context, this paper proposes a model where the ruler can also invest in extractive and productive capacities. In this model, an exogenous or endogenous rise in land productivity can push the economy from a steady state with low extractive capacity, low extraction rate, and low population and output, to a steady state with high extractive capacity, high extraction rate, and high population and output. In that sense, the model can explain how rising land productivity can lead to state building. This might be a more workable interpretation of the Surplus Theory.

References

[1] Acemoglu, D., 2005, Politics and economics in weak and strong states, Journal of Monetary Economics 52, 1199–1226.

[2] Allen, R.C., 1997, Agriculture and the origins of the state in ancient Egypt, Explorations in Economic History 34, 135-154.

[3] Ashraf, Q., and O. Galor, 2011, Dynamics and stagnation in the Malthusian epoch, American Economic Review 101, 2003-2041.

[4] Besley, T., and T. Persson, 2009, The origins of state capacity: property rights, taxation, and politics, American Economic Review 99, 1218-1244.

[5] Besley, T., and T. Persson, 2011, Pillars of Prosperity: The Political Economics of Development Clusters, Princeton University Press, Princeton, New Jersey.

[6] Bockstette, V., A. Chanda, and L. Putterman, 2002, States and markets: the advantage of an early start, Journal of Economic Growth 7, 347–69.

[7] Borcan, O., O. Olsson, and L. Putterman, 2018, State history and economic development: evidence from six millennia, Journal of Economic Growth 23, 1–40.

[8] Chanda, A., and L. Putterman, 2007. Early starts, reversals and catch-up in the process of economic development, Scandinavian Journal of Economics 109, 387–413.

[9] Chanda, A., C.J. Cook, and L. Putterman, 2014, Persistence of fortune: accounting for population movements, there was no post-Columbian reversal, American Economic Journal: Macroeconomics 6, 1-28.

[10] Childe, V.G., 1936, Man Makes Himself, The New American Library, Inc., New York, New York. (Reprinted 1951.)

[11] Dal Bó, E., P., Hernández, S. Mazzuca, 2016, The paradox of civilization: preinstitutional sources of security and prosperity, mimeo, UC Berkeley, NYU Abu Dabi, and Johns Hopkins.

[12] Depetris-Chauvin, E., 2016, State history and contemporary conflict: evidence from Sub-Saharan Africa, mimeo, Pontificia Universidad Católica de Chile.

[13] Depetris-Chauvin, E., and Ö. Özak, 2016, Population diversity, division of labor and comparative development, mimeo, Pontificia Universidad Católica de Chile and Southern Methodist University.

[14] Diamond, J., 1997, Guns, Germs, and Steel: The Fates of Human Societies, W.W. Norton & Company, New York.

[15] Fenske, J. 2014, Ecology, trade, and states in pre-colonial Africa, Journal of the European Economic Association 12, 612–640.

[16] Galor, O., 2010, The 2008 Lawrence R. Klein lecture—comparative economic development: insights from Unified Growth Theory, International Economic Review 51, 1-44.

[17] Hariri, J.G., 2012, The autocratic legacy of early statehood, American Political Science Review 106, 471-494.

[18] Hibbs, D.R. Jr., and O. Olsson, 2004, Geography, biogeography, and why some countries are rich and others are poor, Proceedings of the National Academy of Sciences 101, 3715-3720.

[19] Lagerlöf, N.-P., 2016, Statehood, democracy and preindustrial development, Journal of Economic Dynamics and Control 67, 58-72.

[20] Mann, M., 1986, The Sources of Social Power, Volume 1: A History of Power from the Beginning to AD 1760, 2nd edition, Cambridge University Press, Cambridge.

[21] Mayshar, J., O. Moav, and Z. Neeman, 2017, Geography, transparency, and institutions, American Political Science Review 111, 622–636.

[22] Mayshar, J., O. Moav, Z. Neeman, L. Pascali, 2018, The emergence of hierarchies and states: productivity vs. appropriability, mimeo, Hebrew University of Jerusalem, University of Warwick, Tel-Aviv University, and Pompeu Fabra University.

[23] Nissen, H.J., and P. Heine, 2009, From Mesopotamia to Iraq: A Concise History, The University of Chicago Press, Chicago and London.

[24] Putterman, L., 2008, Agriculture, diffusion and development: ripple effects of the Neolithic Revolution, Economica 75, 729-748.

[25] Read, K.E., 1959, Leadership and consensus in a New Guinea society, American Anthropologist 61, 425-436.

[26] Service, E.R., 1975, Origins of the State and Civilization: The Process of Cultural Evolution, W.W. Norton & Co., New York and London.

[27] Wittfogel, K. A., 1957, Oriental Despotism: A Comparative Study of Total Power, Yale University Press, New Haven.

APPENDIX

A Finding optimal A_{t+1}, z_{t+1} and τ_t

First note from (13) and (14) that output in period $t+1$ can be written

$$Y_{t+1} = (B_{t+1}A_{t+1})^{\alpha}\left[\gamma(1-\tau_t)Y_t\right]^{1-\alpha}. \tag{A1}$$

Substituting $z_{t+1} = \underline{z} + \phi x_t$, (16), and (A1) into (17), we can write U_t^R as a function of A_{t+1}, x_t, and τ_t, namely

$$\begin{aligned} U_t^R = \;& (1-\beta)\ln\left(\tau_t z_t Y_t - \eta A_{t+1}^{\sigma} - x_t\right) \\ & +\beta\ln(\underline{z}+\phi x_t) \\ & +\alpha\beta\ln\left(A_{t+1}\right) \\ & +\beta(1-\alpha)\ln(1-\tau_t) \\ & +\Omega_t \end{aligned} \tag{A2}$$

where

$$\Omega_t = \alpha\beta\ln(B_{t+1}) + \beta(1-\alpha)\ln(Y_t) + \beta(1-\alpha)\ln\gamma$$

contains only variables taken as given by the ruler. The problem is to maximize (A2) subject to $A_{t+1} \geq 0$, $\tau_t \geq 0$, $\tau_t \leq 1$, $x_t \geq 0$, and $x_t \leq (\overline{z}-\underline{z})/\phi$; the last two constraints correspond to $z_{t+1} \geq \underline{z}$ and $z_{t+1} \leq \overline{z}$, respectively.

The first-order conditions for an interior solution state that A_{t+1} and τ_t satisfy

$$(1-\beta)\left[c_t^R\right]^{-1}\eta\sigma A_{t+1}^{\sigma-1} = \alpha\beta\left[A_{t+1}\right]^{-1}, \tag{A3}$$

and

$$(1-\beta)\left[c_t^R\right]^{-1} z_t Y_t = \beta(1-\alpha)\left[1-\tau_t\right]^{-1}, \tag{A4}$$

where $c_t^R = \tau_t z_t Y_t - \eta A_{t+1}^{\sigma} - x_t$; recall (16).

It is straightforward to see that the constraints $A_{t+1} \geq 0$, $\tau_t \geq 0$, and $\tau_t \leq 1$ never bind, so (A3) by (A4) always give optimal A_{t+1} and τ_t for any $x_t \in [0, (\overline{z}-\underline{z})/\phi]$. Using (16), (A3), and (A4) we can solve for ηA_{t+1}^{σ} and $1-\tau_t$ as follows:

$$\eta A_{t+1}^{\sigma} = \left[\frac{\alpha\beta}{\sigma(1-\alpha\beta)+\alpha\beta}\right](z_t Y_t - x_t), \tag{A5}$$

$$1-\tau_t = \left[\frac{\beta\sigma(1-\alpha)}{\sigma(1-\alpha\beta)+\alpha\beta}\right]\left(1-\frac{x_t}{z_t Y_t}\right). \tag{A6}$$

Also, using (16), (A5), and (A6) we can write the ruler's consumption as

$$c_t^R = \left[\frac{\sigma(1-\beta)}{\sigma(1-\alpha\beta)+\alpha\beta}\right](z_t Y_t - x_t). \tag{A7}$$

Below we use (A5) to (A7) to find the optimal choices of A_{t+1} and τ_t for three cases: when $x_t = 0$; when $x_t = (\overline{z}-\underline{z})/\phi$; and when $0 < x_t < (\overline{z}-\underline{z})/\phi$.

A.1 Corner solutions where $x_t = 0$

If the marginal effect on U_t^R from an increase in x_t is negative when $x_t = 0$, then $x_t = 0$ is optimal. This happens when

$$\left.\frac{\partial U_t^R}{\partial x_t}\right|_{x_t=0} = -(1-\beta)\left[\tau_t z_t Y_t - \eta A_{t+1}^{\sigma}\right]^{-1} + \phi\beta\left[\underline{z}\right]^{-1} < 0. \tag{A8}$$

Using (A5) and (A6) we see that $\tau_t z_t Y_t - \eta A_{t+1}^{\sigma}$ is simply the expression for c_t^R in (A7), evaluated at $x_t = 0$. Thus, the inequality in (A8) can be written

$$(1-\beta)\left(\left[\frac{\sigma(1-\beta)}{\sigma(1-\alpha\beta)+\alpha\beta}\right] z_t Y_t\right)^{-1} > \phi\beta\underline{z}^{-1}$$

which translates to $z_t Y_t < \underline{X}$, where $\underline{X}$ is given by (19).

It thus follows that if $z_t Y_t < \underline{X}$, then $x_t = 0$. Moreover, optimal A_{t+1} and τ_t can be found by setting $x_t = 0$ in (A5) and (A6). This gives the bottom rows of (20), and (A16) below.

A.2 Corner solutions where $x_t = (\overline{z} - \underline{z})/\phi$

If the marginal effect on U_t^R from an increase in x_t is positive when $x_t = (\overline{z} - \underline{z})/\phi$, then $x_t = (\overline{z} - \underline{z})/\phi$ is optimal. This happens when

$$\left.\frac{\partial U_t^R}{\partial x_t}\right|_{x_t=\frac{\overline{z}-\underline{z}}{\phi}} = -(1-\beta)\left[\tau_t z_t Y_t - \eta A_{t+1}^{\sigma} - \left(\frac{\overline{z}-\underline{z}}{\phi}\right)\right]^{-1} + \phi\beta\overline{z}^{-1} > 0. \tag{A9}$$

The expression in square brackets in (A9) equals c_t^R in (A7), evaluated at $x_t = (\overline{z} - \underline{z})/\phi$. Substituted into (A9), this gives

$$(1-\beta)\left[\left(\frac{\sigma(1-\beta)}{\sigma(1-\alpha\beta)+\alpha\beta}\right)\left(z_t Y_t - \frac{\overline{z}-\underline{z}}{\phi}\right)\right]^{-1} < \phi\beta\overline{z}^{-1}, \tag{A10}$$

or

$$\overline{z} < \left(\frac{\sigma\beta}{\sigma(1-\alpha\beta)+\alpha\beta}\right)\left[\phi z_t Y_t - (\overline{z}-\underline{z})\right], \tag{A11}$$

which can in turn be simplified to $z_t Y_t > \overline{X}$, where $\overline{X}$ is given by (18).

To sum up, if $z_t Y_t > \overline{X}$, then $x_t = (\overline{z} - \underline{z})/\phi$ and optimal A_{t+1} and τ_t can be found by setting $x_t = (\overline{z} - \underline{z})/\phi$ in (A5) and (A6). This gives the top rows of (20), and (A16) below.

A.3 Interior solutions

Consider next interior solutions for x_t, which can be found when $z_t Y_t \in (\underline{X}, \overline{X})$, and are derived from the first-order condition

$$(1-\beta)\left[c_t^R\right]^{-1} = \phi\beta\left[\underline{z} + \phi x_t\right]^{-1}, \tag{A12}$$

where c_t^R is given by (A7). We can use (A7) and (A12) to find optimal x_t, but we are rather interested in the associated expression for z_{t+1}. Since $z_{t+1} = \underline{z} + \phi x_t$ in an interior solution for x_t, we can write the first-order condition in (A12) as

$$(1-\beta)\left(\left[\frac{\sigma(1-\beta)}{\sigma(1-\alpha\beta)+\alpha\beta}\right]\left[z_t Y_t - \left(\frac{z_{t+1}-\underline{z}}{\phi}\right)\right]\right)^{-1} = \phi\beta\left[z_{t+1}\right]^{-1}. \tag{A13}$$

This can be solved for z_{t+1} to give

$$z_{t+1} = \underline{z} + \phi x_t = \left(\frac{\beta\sigma}{\beta\sigma(1-\alpha)+\sigma+\alpha\beta}\right)\left[\phi z_t Y_t + \underline{z}\right], \tag{A14}$$

which is the middle row of (21).

Using (A14), we can also derive the associated solutions for τ_t and A_{t+1}. Dividing (A3) by (A12), and rearranging, investment in next period's technology becomes

$$\eta A_{t+1}^{\sigma} = \frac{\alpha}{\sigma\phi}\left(\underline{z}+\phi x_t\right) = \left(\frac{\alpha\beta}{\beta\sigma(1-\alpha)+\sigma+\alpha\beta}\right)\left(z_t Y_t + \frac{\underline{z}}{\phi}\right), \tag{A15}$$

where the second equality follows from (A14). Solving (A15) for A_{t+1} gives the middle row in (A16) below.

Similarly, dividing (A4) by (A12), and rearranging, gives

$$(1-\tau_t)z_t Y_t = \left(\frac{1-\alpha}{\phi}\right)\left[\underline{z}+\phi x_t\right] = \left(\frac{\beta\sigma(1-\alpha)}{\beta\sigma(1-\alpha)+\sigma+\alpha\beta}\right)\left(z_t Y_t + \frac{\underline{z}}{\phi}\right)$$

which can be solved to give the middle row in (20).

A.4 Complete characterization of the solution

To sum up, we can write the optimal expressions for τ_t as in (20), for z_{t+1} as in (21), and the optimal expression for A_{t+1} can be written

$$A_{t+1} = \begin{cases} \left[\frac{1}{\eta}\left(\frac{\alpha\beta}{\sigma(1-\alpha\beta)+\alpha\beta}\right)\left[z_t Y_t - \left(\frac{\overline{z}-\underline{z}}{\phi}\right)\right]\right]^{\frac{1}{\sigma}} & \text{if} \quad z_t Y_t \geq \overline{X}, \\ \left[\frac{1}{\eta}\left(\frac{\alpha\beta}{\beta\sigma(1-\alpha)+\sigma+\alpha\beta}\right)\left(z_t Y_t + \frac{\underline{z}}{\phi}\right)\right]^{\frac{1}{\sigma}} & \text{if} \quad z_t Y_t \in \left[\underline{X}, \overline{X}\right], \\ \left[\frac{1}{\eta}\left(\frac{\alpha\beta}{\sigma(1-\alpha\beta)+\alpha\beta}\right) z_t Y_t\right]^{\frac{1}{\sigma}} & \text{if} \quad z_t Y_t \leq \underline{X}. \end{cases} \tag{A16}$$

B Dynamics of Y_t

This section finds an expression for Y_{t+1} in terms of Y_t, z_t, and B_{t+1}. Consider first the case when $z_t Y_t \in \left[\underline{X}, \overline{X}\right]$, meaning neither of the constraints on z_{t+1} binds. Using (20), it is then seen that

$$\gamma(1-\tau_t)Y_t = \frac{\gamma}{\phi z_t}\left[\frac{\beta\sigma(1-\alpha)}{\beta\sigma(1-\alpha)+\sigma+\alpha\beta}\right](\phi z_t Y_t + \underline{z}). \tag{A17}$$

Now (A1), (A16), and (A17) tell us that

$$\begin{aligned} Y_{t+1} = B_{t+1}^{\alpha} & \underbrace{\left(\frac{1}{\eta\phi}\right)^{\frac{\alpha}{\sigma}}\left(\frac{\alpha\beta}{\beta\sigma(1-\alpha)+\sigma+\alpha\beta}\right)^{\frac{\alpha}{\sigma}}(\phi z_t Y_t + \underline{z})^{\frac{\alpha}{\sigma}}}_{A_{t+1}^{\alpha}} \\ & \times \underbrace{\left(\frac{\gamma}{\phi z_t}\right)^{1-\alpha}\left(\frac{\beta\sigma(1-\alpha)}{\beta\sigma(1-\alpha)+\sigma+\alpha\beta}\right)^{1-\alpha}[\phi z_t Y_t + \underline{z}]^{1-\alpha}}_{[\gamma(1-\tau_t)Y_t]^{1-\alpha}}. \end{aligned} \tag{A18}$$

To simplify this expression, first define

$$\rho = \frac{\alpha}{\sigma} + 1 - \alpha \in (0,1) \tag{A19}$$

and

$$D = \left[\frac{\alpha}{\eta\sigma(1-\alpha)}\right]^{\frac{\alpha}{\sigma}}\gamma^{1-\alpha}\left(\frac{1}{\phi}\right)^{\rho}\left[\frac{\beta\sigma(1-\alpha)}{\beta\sigma(1-\alpha)+\sigma+\alpha\beta}\right]^{\rho}, \tag{A20}$$

where we note that $\rho < 1$ follows from $\sigma > 1$. Using (A19) and (A20), we can rewrite (A18) more compactly as

$$Y_{t+1} = DB_{t+1}^{\alpha} z_t^{\alpha-1}(\phi z_t Y_t + \underline{z})^{\rho}, \tag{A21}$$

which is the middle row of (22).

Consider next the case when $z_t Y_t > \overline{X}$. From (20), it now follows that

$$\gamma(1-\tau_t)Y_t = \frac{\gamma}{\phi z_t}\left[\frac{\beta\sigma(1-\alpha)}{\sigma(1-\alpha\beta)+\alpha\beta}\right](\phi z_t Y_t + \underline{z} - \overline{z}). \tag{A22}$$

Following similar steps as we followed above for the interior solution, we can use (A1), (A16), and (A22) to show that

$$Y_{t+1} = \widehat{D}B_{t+1}^{\alpha} z_t^{\alpha-1}(\phi z_t Y_t + \underline{z} - \overline{z})^{\rho}, \tag{A23}$$

where

$$\widehat{D} = \left[\frac{\alpha}{\eta\sigma(1-\alpha)}\right]^{\frac{\alpha}{\sigma}}\gamma^{1-\alpha}\left(\frac{1}{\phi}\right)^{\rho}\left[\frac{\beta\sigma(1-\alpha)}{\beta\sigma(1-\alpha)+(1-\beta)\sigma+\alpha\beta}\right]^{\rho}, \tag{A24}$$

and (recall) ρ is given by (A19).

Finally, for the case when $z_t Y_t < \underline{X}$, we can use (20) again to see that

$$\gamma(1-\tau_t)Y_t = \gamma\left[\frac{\beta\sigma(1-\alpha)}{\beta\sigma(1-\alpha)+(1-\beta)\sigma+\alpha\beta}\right]Y_t. \tag{A25}$$

Applying (A1), (A16), and (A25) some algebra shows that

$$Y_{t+1} = \widehat{D} B_{t+1}^{\alpha} \left(\phi z_t Y_t\right)^{\rho}, \tag{A26}$$

where ρ and $\widehat{D}$ are given by (A19) and (A24).

Finally, using (A19), (A20), and (A24) we can define κ as

$$\kappa = \frac{\widehat{D}}{D} = \left[\frac{\beta\sigma(1-\alpha)+\sigma+\alpha\beta}{\beta\sigma(1-\alpha)+(1-\beta)\sigma+\alpha\beta}\right]^{\rho} > 1. \tag{A27}$$

Now, using (A23), (A26), and substituting for $\widehat{D} = \kappa D$, we arrive at the top and bottom rows of (22).

C The phase diagram

C.1 The $(z_{t+1} = z_t)$-locus

The following can be seen directly from (21): for $z_t Y_t \le \underline{X}$, it holds that $z_{t+1} = z_t$ when $z_t = \underline{z}$; for $z_t Y_t \ge \overline{X}$, it holds that $z_{t+1} = z_t$ when $z_t = \overline{z}$; for $z_t Y_t \in \left[\underline{X}, \overline{X}\right]$, it holds that $z_{t+1} = z_t$ when $z_t = \left(\frac{\beta\sigma}{\beta\sigma(1-\alpha)+\sigma+\alpha\beta}\right)\left[\phi z_t Y_t + \underline{z}\right]$, or $z_t = \beta\sigma\underline{z}/\{\beta\sigma(1-\alpha)+\sigma+\alpha\beta-\beta\sigma\phi Y_t\}$. In sum, the $(z_{t+1} = z_t)$-locus can be written

$$z_t = \begin{cases} \overline{z} & \text{if} \quad z_t Y_t \ge \overline{X}, \\ \frac{\beta\sigma\underline{z}}{\beta\sigma(1-\alpha)+\sigma+\alpha\beta-\beta\sigma\phi Y_t} & \text{if} \quad z_t Y_t \in \left[\underline{X}, \overline{X}\right], \\ \underline{z} & \text{if} \quad z_t Y_t \le \underline{X}. \end{cases} \tag{A28}$$

The (inverse of) (A28) is graphed in Figure 2 as a three-segment solid blue curve.

C.2 The $(Y_{t+1} = Y_t)$-locus

Setting $B_{t+1} = B_0$ in (22) we learn the following: for $z_t Y_t \le \underline{X}$, it holds that $Y_{t+1} = Y_t$ when $Y_t = \left[\kappa D B_0^{\alpha}\left(\phi z_t\right)^{\rho}\right]^{1/(1-\rho)}$; for $z_t Y_t \ge \overline{X}$, it holds that $Y_{t+1} = Y_t$ when $Y_t = \xi(z_t, B_0)$, defined from $\xi(z_t, B_0) = \kappa D B_0^{\alpha} z_t^{\alpha-1}\left[\phi z_t \xi(z_t, B_0) + \underline{z} - \overline{z}\right]^{\rho}$; for $z_t Y_t \in \left[\underline{X}, \overline{X}\right]$, it holds that $Y_{t+1} = Y_t$ when $Y_t = \vartheta(z_t, B_0)$, defined from $\vartheta(z_t, B_0) = \kappa D B_0^{\alpha} z_t^{\alpha-1}\left[\phi z_t \vartheta(z_t, B_0) + \underline{z}\right]^{\rho}$. To summarize, the $(Y_{t+1} = Y_t)$-locus can be written

$$Y_t = \begin{cases} \xi(z_t, B_0) & \text{if} \quad z_t Y_t \ge \overline{X}, \\ \vartheta(z_t, B_0) & \text{if} \quad z_t Y_t \in \left[\underline{X}, \overline{X}\right], \\ \left[\kappa D B_0^{\alpha}\left(\phi z_t\right)^{\rho}\right]^{\frac{1}{1-\rho}} & \text{if} \quad z_t Y_t \le \underline{X}. \end{cases} \tag{A29}$$

The red solid curves in Figure 2 show the graphs of the three different segments of the $(Y_{t+1} = Y_t)$-locus in (A29).

C.3 Change in configuration when changing B_0

Note that the $(z_{t+1} = z_t)$-locus does not depend on B_0. It is easy to see, from the definitions above, that $\xi(z_t, B_0)$ and $\vartheta(z_t, B_0)$ are strictly increasing in B_0, that $\lim_{B\to\infty} \xi(z_t, B) = \lim_{B\to\infty} \vartheta(z_t, B) = \infty$, and that $\xi(z_t, 0) = \vartheta(z_t, 0) = 0$. It follows that we can adjust B_0 to shift the $(Y_{t+1} = Y_t)$-locus to alter the configuration of the two-dimensional dynamical system. When B_0 is sufficiently small the $(Y_{t+1} = Y_t)$- and $(z_{t+1} = z_t)$-loci intersect only once, and this unique intersection lies in the region where $z_t Y_t < \underline{X}$. When B_0 is sufficiently large the two loci also intersect only once, now in the region where $z_t Y_t > \overline{X}$.

D Proof of Proposition 2

(a) Let $\widehat{B}$ be defined as the level of B_0 that generates $\underline{zY} = \underline{X}$. From (24) follows that

$$\widehat{B} = \left[\frac{\underline{X}^{1-\rho}}{\phi^{\rho}\kappa D \underline{z}}\right]^{\frac{1}{\alpha}} > 0, \tag{A30}$$

where $\underline{X}$ is given by (19). Note from (24) that $\underline{Y}$ is increasing in B_0. That is, if $B_0 < \widehat{B}$, then $\underline{zY} < \underline{X}$.

(b) Let $\widehat{\widehat{B}}$ be defined as the level of B_0 that generates $\overline{z}\overline{Y} = \overline{X}$. Recall that imposing steady state in the top row of (22) defines $\overline{Y}$ from $\overline{Y} = \kappa D B_0^{\alpha} \overline{z}^{\alpha-1} \left[\phi \overline{z}\overline{Y} + \underline{z} - \overline{z}\right]^{\rho}$, which shows that $\overline{Y}$ is increasing in B_0. Setting $B_0 = \widehat{\widehat{B}}$ and $\overline{z}\overline{Y} = \overline{X}$ gives

$$\widehat{\widehat{B}} = \frac{1}{\overline{z}}\left[\frac{\overline{X}}{\kappa D\left(\phi\overline{X} + \underline{z} - \overline{z}\right)^{\rho}}\right]^{\frac{1}{\alpha}} = \frac{1}{\overline{z}}\left[\frac{\overline{X}}{\kappa D\left[\phi\underline{X}\left(\frac{\overline{z}}{\underline{z}}\right)\right]^{\rho}}\right]^{\frac{1}{\alpha}} > 0, \tag{A31}$$

where the second equality follows from (18) and (19). Since $\overline{Y}$ is increasing in B_0, it follows that $B_0 > \widehat{\widehat{B}}$ implies $\overline{z}\overline{Y} > \overline{X}$.

(c) Using (A30) and (A31), and letting things cancel, we can write $\widehat{\widehat{B}} < \widehat{B}$ as

$$\frac{\underline{X}}{\overline{X}} < \overline{z}^{-(\alpha+\rho)}\underline{z}^{1+\rho}. \tag{A32}$$

Next, using (18) and (19), and dividing both sides by $\underline{z}$, we can write the inequality in (A32) as

$$\frac{\sigma(1-\alpha\beta) + \alpha\beta}{\overline{z}\left[\beta\sigma(1-\alpha) + \sigma + \alpha\beta\right] - \beta\sigma\underline{z}} > \overline{z}^{-(\alpha+\rho)}\underline{z}^{\rho}. \tag{A33}$$

The inequality in A33 thus implies $\widehat{\widehat{B}} < \widehat{B}$. Letting $\underline{z}$ go to zero, the left-hand side of (A33) goes to something strictly positive, while the right-hand side goes to zero; recall that $\rho > 0$.

Thus, the inequality in (A33) must hold for $\underline{z}$ sufficiently close to zero, implying in turn that $\widehat{\widehat{B}} < \widehat{B}$ holds for $\underline{z}$ sufficiently close to zero.

(d)

Part (i): From parts (a) and (b) of this Proposition, we know that the two steady states are constructed such that $\overline{Y} > \overline{X}/\overline{z}$, and $\underline{Y} < \underline{X}/\underline{z}$. From (18) and (19) it follows that

$$\frac{\overline{X}}{\overline{z}} - \left(\frac{\overline{z} - \underline{z}}{\phi}\right) = \frac{\underline{X}}{\underline{z}}, \tag{A34}$$

implying that $\overline{Y} > \overline{X}/\overline{z} > \underline{X}/\underline{z} > \underline{Y}$ (since $\overline{z} > \underline{z}$).

Part (ii): Using (14) the population levels in the two steady states can be written $\overline{L} = \gamma(1-\overline{\tau})\overline{Y}$ and $\underline{L} = \gamma(1-\underline{\tau})\underline{Y}$, respectively. From (A6) follows that

$$\overline{L} = \gamma(1-\overline{\tau})\overline{Y} > \gamma\left[\frac{\beta\sigma(1-\alpha)}{\sigma(1-\alpha\beta)+\alpha\beta}\right]\left[\frac{\overline{X}}{\overline{z}} - \left(\frac{\overline{z} - \underline{z}}{\phi}\right)\right], \tag{A35}$$

where we have used $x_t = (\overline{z} - \underline{z})/\phi$ and $\overline{Y} > \overline{X}/\overline{z}$; recall that $x_t = (\overline{z} - \underline{z})/\phi$ when $\overline{z}\overline{Y} > \overline{X}$, i.e., when $z_{t+1} \leq \overline{z}$ binds. Using (A6) again, we see that

$$\underline{L} = \gamma(1-\underline{\tau})\underline{Y} < \gamma\left[\frac{\beta\sigma(1-\alpha)}{\sigma(1-\alpha\beta)+\alpha\beta}\right]\frac{\underline{X}}{\underline{z}}, \tag{A36}$$

where we have used $x_t = 0$ and $\underline{Y} < \underline{X}/\underline{z}$; recall that $x_t = 0$ when $\underline{Y} < \underline{X}/\underline{z}$, i.e., when $z_{t+1} \geq \underline{z}$ binds. Now (A34), (A35), and (A36) together imply that $\overline{L} > \underline{L}$.

Part (iii): The result follows from (20). The bottom row equals $\underline{\tau} = [\sigma(1-\beta)+\alpha\beta]/[\sigma(1-\alpha\beta)+\alpha\beta]$, and $\overline{\tau}$ is defined as the top row, evaluated at $z_t Y_t = \overline{z}\overline{Y}$, which can be written:

$$\overline{\tau} = 1 - (1-\underline{\tau})\left[1 - \left(\frac{\overline{z} - \underline{z}}{\phi}\right)\frac{1}{\overline{z}\overline{Y}}\right] = \underline{\tau} + \left(\frac{\overline{z} - \underline{z}}{\phi}\right)\frac{1-\underline{\tau}}{\overline{z}\overline{Y}} > \underline{\tau}. \tag{A37}$$

∎

Figures

Figure 1: Illustration of the workings of Model A. The bottom panel shows the dynamics of y_t, and the top panel the determination of τ_t in any given period.

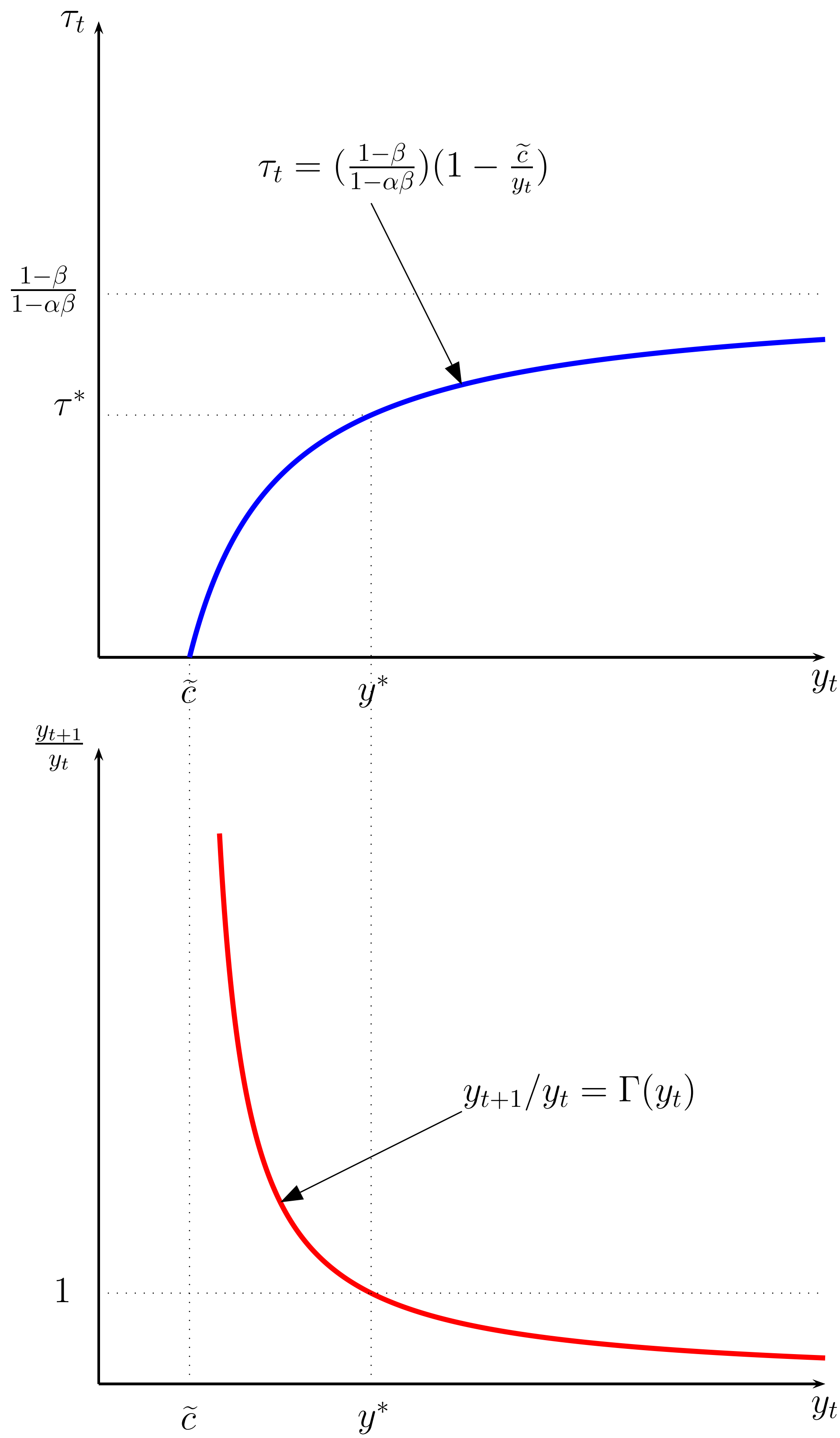

Figure 2: Phase diagram illustrating the dynamics of Model B, when setting $B_t = B_0$ for all t. The loci along which z_t and Y_t are constant are indicated by the red and blue solid curves. The green dashed curves indicate the loci above and below which the constraints $z_{t+1} \leq \overline{z}$ and $z_{t+1} \geq \underline{z}$ bind. In this configuration, there exist two stable steady states.

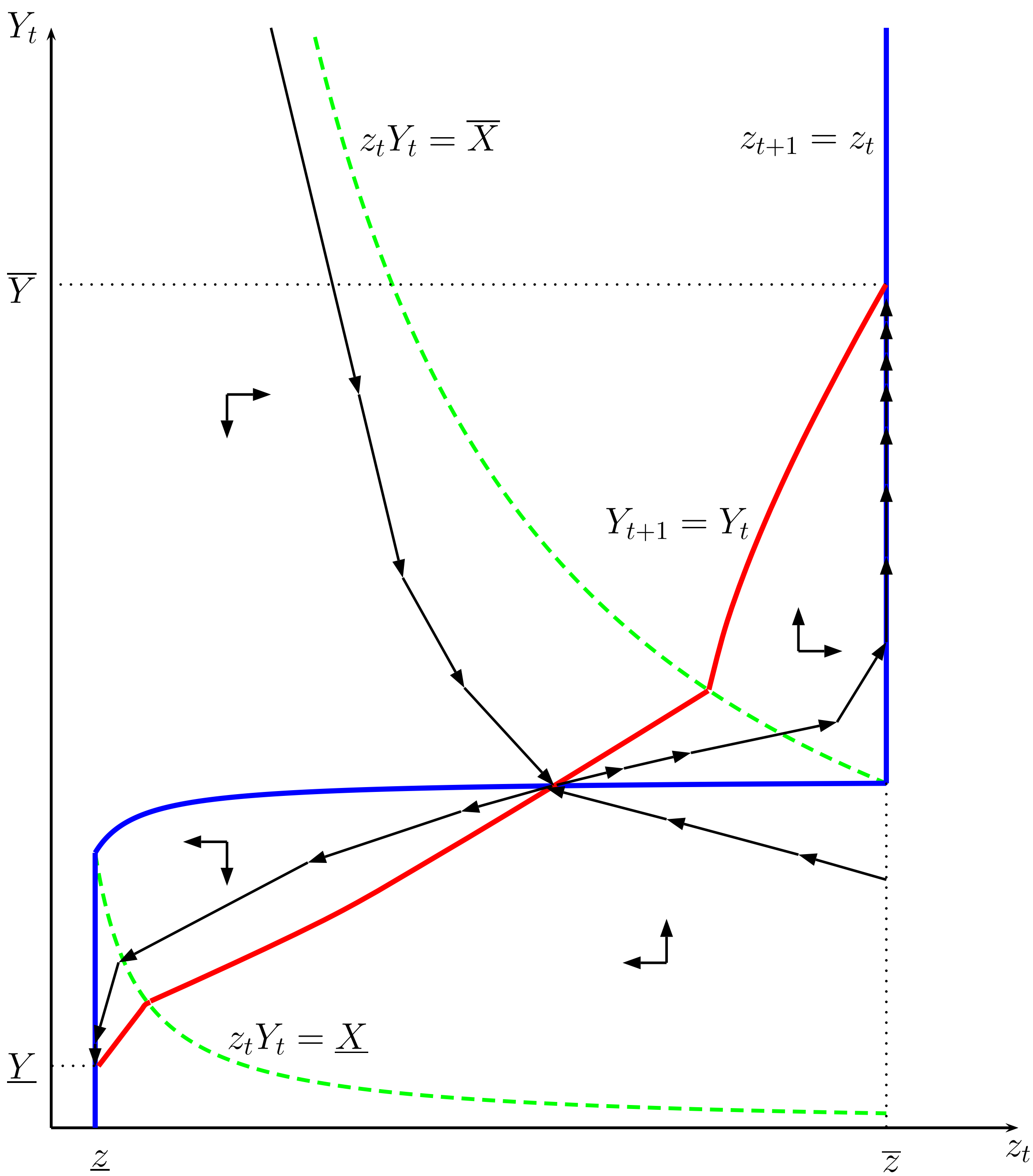

Figure 3: Simulated time paths when B_t evolves endogenously.