Submission Number: PET11-11-00036

The Grand Experiment of Communism: Discovering the Trade-off between Equality and Efficiency

Alireza Naghavi
University of Bologna

Alexander Mihailov
University of Reading

Etienne Farvaque
Universite de Lille 1

Abstract

The rise and fall of communism is a complex theme examined from different theoretical and methodological perspectives in social sciences. Our approach aims to formalize the interactions of economic incentives and social learning through experimentation with an unattempted economic system. We introduce inequality-averse and inefficiency-averse agents responding to incentives and transmitting their ideology as they are affected by evolving outcomes. We analyze their conflict through a repeated game between a leader with economic power and followers with ideological determination. The socioeconomic dynamics of our model generate a pendulum-like switch from markets to a centrally-planned economy abolishing private ownership, and back to restoring market incentives. The grand experiment of communism is thus characterized to have led to the discovery of a trade-off between equality and efficiency at the scale of alternative economic systems.

Submitted: February 15, 2011.

The Grand Experiment of Communism: Discovering the Trade-off between Equality and Efficiency

Etienne Farvaque* Alexander Mihailov† Alireza Naghavi‡

February 2011§

Abstract

The rise and fall of communism is a complex theme examined from different theoretical and methodological perspectives in social sciences. Our approach aims to formalize the interactions of economic incentives and social learning through experimentation with an unattempted economic system. We introduce inequality-averse and inefficiency-averse agents responding to incentives and transmitting their ideology as they are affected by evolving outcomes. We analyze their conflict through a repeated game between a leader with economic power and followers with ideological determination. The socioeconomic dynamics of our model generate a pendulum-like switch from markets to a centrally-planned economy abolishing private ownership, and back to restoring market incentives. The grand experiment of communism is thus characterized to have led to the discovery of a trade-off between equality and efficiency at the scale of alternative economic systems.

Key words: capitalism, communism, inequality, inefficiency, transmission of beliefs, economic transitions

JEL classification codes: C72, D31, D63, D74, D83, P51

*Equippe – Universités de Lille, Faculté des Sciences Economiques et Sociales, Universite de Lille 1, 59655 Villeneuve d'Ascq Cedex France; Etienne.Farvaque@univ-lille1.fr.

†Department of Economics, University of Reading, Whiteknights, Reading RG6 6AA, and Department of Economics, University of Warwick, Coventry, CV4 7AL, United Kingdom; a.mihailov@reading.ac.uk.

‡Department of Economics, University of Bologna, Piazza Scaravilli 2, Bologna 40126, Italy. alireza.naghavi@unibo.it

§First draft: January 2011. Mihailov acknowledges funding from the University of Reading Board for Research in the Social Sciences (Project E3310100, 2010-2011). Mihailov and Naghavi thank University of Lille 1 for granting each of them an invited professorship that greatly facilitated the work along the project. We are also grateful to our Universities for the hospitality while working together at various stages and to many colleagues for stimulating discussions. In particular, we would like to thank for detailed feedback Mark Casson, Rumen Dobrinsky, Nigar Hashimzade, Gaël Lagadec, and seminar participants at Reading (February 2011). The responsibility for any errors and misinterpretaions remains our own.

Contents

List of Figures

> February 1848: "The Communists disdain to conceal their views and aims. They openly declare that their ends can be attained only by the forcible overthrow of all existing social conditions. Let the ruling classes tremble at a Communistic revolution. The proletarians have nothing to lose but their chains. They have a world to win." Karl Marx and Frederick Engels, *Manifesto of the Communist Party.*[1]

> 8 September 2010: "There were many odd things about my recent Havana stopover [...] but one of the most unusual was Fidel Castro's level of self-reflection. [...] I asked him if he believed the Cuban model was still something worth exporting. 'The Cuban model doesn't even work for us anymore,' he said." Jeffrey Goldberg, "Fidel: 'Cuban Model Doesn't Even Work for Us Anymore'," *The Atlantic.*[2]

1 Why Another Fable of the Grand Experiment?

The grand illusion of the 20-th century was, no doubt, communism. It was also its grand experiment. In terms of utopian vision, radical implementation and socioeconomic impact, communism has left a lasting mark in history.

The rise and fall of communism is a complex theme interpreted from different theoretical and methodological perspectives in social sciences. It appears to be a phenomenon so fundamental and multi-faceted that its study and comprehensive understanding has warranted joining forces from many disciplines going beyond politics and economics and into the domains of sociology, philosophy, culture and the arts. Not surprisingly, millions of articles, books and films, both scientific and popular, have addressed the grand experiment of the 20-th century. These have tried to portray or, more ambitiously, explain the various manifestations of communism across the map of the world – from nascent and militant through mature and oppressive into stagnating and decaying.

So why another attempt to reconsider the key driving mechanisms behind the genesis of the revolutionary communist project and the gradual mass disillusionment with its realities? The novelty of our approach consists in using economic theory to examine the interactive dynamics of economic incentives and social learning through experimentation with an economic system that has never been implemented before. Our interest, and focus, here is not in the scenario already analyzed by Acemoglu and Robinson (2000), where democratic reforms such as the extension of the franchise and the increase of redistribution have enabled Western European countries to avoid the revolutionary advent of communism. We, instead, look at the alternative scenario that led to the October Revolution of 1917 and the subsequent establishment of the Soviet Union in a big region

[1] Chapter IV. Position of the Communists in Relation to the Various Existing Opposition Parties, translated by Samuel Moore in cooperation with Frederick Engels, 1888, http://www.marxists.org/archive/marx/works/1848/communist-manifesto/ch04.htm

[2] http://www.theatlantic.com/international/archive/2010/09/fidel-cuban-model-doesnt-even-workfor-us-anymore/62602/

of the world where capitalism was less developed and, perhaps more importantly, any democratization of the society was avoided or much delayed, as Lenin argued. Along this second scenario, the determination of the oppressed to unite, coordinate and stage a revolution was an outcome of increasing resentment with the old, 'unfair' system of 'exploitation of the proletariat by the bourgeoisie', catalyzed by the emergence and leadership of a communist party. This is, in essence, the mechanism of the class conflict which brought about the communist experiment in Russia.

In what follows, we formalize the socioeconomic process that led to communism via a forced revolution and nationalization of capital, as well as its reversal back to markets. We believe this to be an important insight from history over the last century and a half: in essence, the grand experiment of communism has led to the discovery of a trade-off between equality and efficiency at the scale of alternative economic systems. Our theoretical account of the rise and fall of communism, from the revolutionary enthusiasm of Marx and Engels through the disillusionment of Castro we quoted in the very beginning, is framed as a stylized game of class struggle involving economic decisions, transmission of ideology across generations and social learning.

We model two types of agents, inequality-averse and inefficiency-averse ones, responding to economic incentives and transmitting their values as they are affected by evolving economic outcomes. Even though communism has been thoroughly studied,[3] we are not aware of any account produced along the lines of the transmission and evolution of the preferences of agents. We deem such a perspective fruitful and not negligible. The importance of the transmission of beliefs has already been explored in other contexts (see, e.g., the survey by Bisin and Verdier, 2010), but not with respect to regime switches across economic systems under a dynamic growth process.[4] In particular, we first show how capital accumulation by the minority elite and the resulting inequality leads to increasing social discontent over time and, eventually, the overthrow of the system. We then show how a centrally-planned system aimed at equality also falls apart due to misalignment of individual and aggregate incentives, lower well-being and the gradual redirection of ideas towards a market system.

Our principal result is to demonstrate how the mutually dependent processes driving the longer-run socioeconomic dynamics of our model can generate such pendulum-like switch from markets under weak democracy to a mono-party system abolishing private ownership and imposing a centrally-planned economy, and back to rebuilding market incentives and political pluralism to sustain society. The economic literature, and the literature on communism or social evolution more generally, has not provided so far a consistent theory on the experience of the Soviet Union and its satellite countries in Eastern Europe throughout the 20-th century accounting for both these transitions. In this consists the originality of our simple and stylized but historically trust-worthy formal analysis of the rise and fall of communism.

The paper is further down organized as follows. In the next section we construct our

[3] Among many others, by Lange (1956 [1936]), von Hayek (1940), Tinbergen (1960), Kornai (1980), Roemer (1985), Brown (2010).

[4] See also Saint-Paul (2010) on the impact of the evolution of beliefs on ideological bias in the society and political reform.

model, presenting the types of agents, their objectives, constraints, ideological conflict, and the transmission of their beliefs across generations. Section 3 then solves the von Stackelberg game between the agent types that drives the intergenerational dynamics and highlights our key analytical results. Section 4 offers a broader discussion, placing our work in the context of the closest literature, and the last section concludes. More details concerning the inherited capital stock carried over the economic transitions are provided in Appendix A, while Appendix B derives the solutions of the within-period von Stackelberg game, itself repeated over time. The figures we refer to in the main text are collected in Appendix C.

2 The Model

We consider two economic systems under which society can evolve: one is market-based (capitalist), denoted by M, and the other centrally-planned (communist), denoted by C. True to the historical genesis of communist ideas, our analysis begins with a market-based system founded upon property rights over the means of production and the corresponding private incentives to capital accumulation. We, in effect, look at a 'global economy' which initially is operating under a market capitalist system, but at some point in time a region splits apart and experiments with communism. We model the driving forces towards communism, and after learning from experience, back to markets. Our interest is in the particular world region that passes through the grand socioeconomic experiment. Accordingly, the rest of the world which remains market-based, and possibly implements gradual democratization and redistribution (e.g., as in Acemoglu and Robinson, 2000), is just a reference point for the agents.

2.1 Agent Types, Economic Power and the Conflict Function

There are two types of agents in the initial capitalist society. The large majority are born without inheriting capital: they are the 'workers'. They cannot extract rents from ownership of the capital stock. In this sense, workers are 'unprivileged' in the capitalist society, and can only sell their labor force in the market in order to subsist, as Marx argued. Being the have-nots, what they care about is inequality in the capitalist society, whose victim they are by birth. We call them *inequality-averse* agents and denote their type as A. A minority of agents are born with inherited capital: they are the 'capitalists'. Each of them extracts rents from his ownership over a proportion of the total capital stock. Capitalists thus have a 'privileged social status', namely, not to have to sell their labor force in the market. That is why they care about inefficiency of production, as it reduces profits. We call them *inefficiency-averse* agents and denote their type as B.

Initially, types (A and B) and 'classes' (workers and capital owners) coincide, by definition. In a conventional way, this can be interpreted in terms of the class struggle between capital and labor. However, over time preference types can evolve, so that class and preference type may diverge.

Under both systems, M and C, *economic power* belongs to the preference type who decides upon – and enforces – the intertemporal allocation between capital accumulation

and consumption. The other preference type can then only try to change the economic system through ideological influence. We denote the degree of *strength* of each type *relative* to the other by the conflict function $q_t(\cdot)$ and $1-q_t(\cdot)$, respectively for types A and B in any period t, and measure it by an index, $0 < q_t(\cdot) < 1$. More precisely, this index can be defined to be some increasing function of the *relative size* (or fraction) of each preference type in the total adult population in the considered world region, $0 < q_{S,t} < 1$, and of the *relative intensity* of the preference itself (or ideological determination), $0 < q_{I,t} < 1$: $0 < q_t(q_{S,t}, q_{I,t}) < 1$, with $\frac{\partial q_t(\cdot)}{\partial q_{S,t}} > 0$ and $\frac{\partial q_t(\cdot)}{\partial q_{I,t}} > 0$.[5] The law of motion of $q_t(\cdot)$ – and, hence, of $1-q_t(\cdot)$ – will turn out to depend on socialization efforts. The conflict function represents the degree of social resentment or ideological determination to change the status quo, $q_t(\cdot)$ for type A and $1-q_t(\cdot)$ for type B. It will also define the probability of a regime shift in any period t. In what follows, we assume a switch from one economic system to another occurs once the strength of the oppressed type dominates that of the ruling type: $q_t(\cdot) > 0.5$ for A and $1-q_t(\cdot) > 0.5$ for B.

[Figure 1 about here]

Figure 1 presents the timing and the structure of the model we develop and analyze in the present paper. As mentioned, we begin at time $t = 0$ with a 'global economy' functioning under a capitalist system. If, at some $t = T$, the relative strength of type A to change the status quo exceeds the critical value, a communist revolution occurs to establish equality by nationalizing the capital stock. Similarly, our model will show that inferior standards of living under communism can induce type A agents to join pro-market thinking even if their predecessors were not capital owners, which will allow for belief convergence to type B. Thus, if at some $t = T + n$, the relative strength of type B to change the status quo exceeds the critical value, privatization of the capital stock arises to re-establish efficiency through market incentives.

2.2 Utility and Preferences

Henceforth, we focus on the region of the world that undergoes the communist experiment. Its total adult population is normalized to 1.

The utility of agent i for $i = A, B$ under each system $j = C, M$ takes the form

$$U^A_{j,t}(c^A_{j,t}, Y^A_{j,t+1}, \upsilon_t) = c^A_{j,t} + \beta Y^A_{j,t+1} - E^A_t(\upsilon_{t+1}) - \frac{\tau^A_t(\upsilon_t)^2}{2}, \quad (1)$$

$$U^B_{j,t}(c^B_{j,t}, Y^B_{j,t+1}, \chi_t) = c^B_{j,t} + \beta Y^B_{j,t+1} - E^B_t(\chi_{t+1}) - \frac{\tau^B_t(\chi_t)^2}{2}, \quad (2)$$

[5] A specific functional form could be, for instance, a weighted average of the two arguments: $q_t(q_{S,t}, q_{I,t}) \equiv \theta_t q_{S,t} + (1-\theta_t) q_{I,t}$, with $0 < \theta_t < 1$ or, simply, $\theta = 0.5$. This decomposition of $q_t(\cdot)$ may be useful in drawing some contrasts in terms of which of the two arguments is operative in the dynamics of the function under markets versus communism, as we shall briefly make clear when appropriate further below. Yet we prefer to keep the generality of the definition as $q_t(\cdot)$, which also preserves the validity for both economic systems of the key dynamic equation we are going to derive and interpret.

with $c^i_{j,t}$ denoting individual consumption levels, $Y^i_{j,t+1}$ returns to potential private investment, and $0 < \beta < 1$ the discount factor identical for all agents. The first two terms in the utility function we specify are standard. The next two are less so, although they too have analogues in the broader literature: the third term generally implies that relative status, or reference points, with respect to others matter as well; and the fourth term captures costly socialization efforts (in the transmission of values across generations). More precisely, $E^A_t(\upsilon_{t+1})$ and $E^B_t(\chi_{t+1})$ are the expected inequality and inefficiency that depend on the regime in the next period

$$E^A_t(\upsilon_{t+1}) = q_{t+1} \ln \upsilon_{C,t+1} + (1 - q_{t+1}) \ln \upsilon_{M,t+1}, \quad (3)$$

$$E^B_t(\chi_{t+1}) = q_{t+1} \ln \chi_{C,t+1} + (1 - q_{t+1}) \ln \chi_{M,t+1}, \quad (4)$$

where $\upsilon_{j,t} = \frac{Y^B_{j,t}}{Y^A_{j,t}}$ measures inequality within the society at t, and $\chi_{j,t}$ the relative efficiency of individual optimization over that by a centrally planned system. Note that communism forcefully proclaims complete equality in the society, $\upsilon_{C,t} = \frac{Y^B_{C,t}}{Y^A_{C,t}} = 1$, yielding $\ln \upsilon_C = 0$. Similarly, efficiency is initially normalized under capitalism, when no other economic system has as yet emerged to compare against: $\chi_{M,t} = 1$, and so $\ln \chi_M = 0$. $\upsilon_{M,t}$ and $\chi_{C,t}$ will be defined further below. As can be seen from (3) and (4), the strength of both preference types in any period t, q_t and $1 - q_t$, determine the probability of a regime change in that period. Finally, utility depends on costly socialization effort functions $\tau^A_t(\upsilon_t)$ and $\tau^B_t(\chi_t)$, with $0 \leq \tau^i_t(\cdot) \leq 1$, to be discussed in section 2.6.

2.3 Production and Incentives

We also employ a relatively standard production function, which can be written as

$$H_{j,t}(A_j, K_t, L) = A_j(c^A_{j,t})[\alpha K_t + (1 - \alpha)L] \quad (5)$$

for $j = C, M$ and depends on two productive factors, labor L supplied infinitely by households and capital K_t. Different from Acemoglu and Robinson (2000) who assume exogenous constant productivity higher in the market relative to home production in a somewhat related context, our formulation in (5) endogenizes the productivity of each regime, $A_j(c^A_{j,t})$, as a function of the incentives of workers to take efficient part in the production process or, in a more direct sense, of the (dis)satisfaction of workers with their material well-being approximated by their consumption level. There is also a technological parameter, α, measuring the contributions to output of capital and labor. Competitive factor markets under capitalism equalize marginal factor products to marginal factor returns, so wages and capital rents can be written as

$$w_{j,t} = (1 - \alpha)A_j(c^A_{j,t})$$

and

$$r_{j,t} = \alpha A_j(c^A_{j,t}).$$

We assume $\alpha > 1/2$ so that returns to capital always exceed returns to labor, that is, no capitalist finds the incentives to become a worker. Note that both factor returns

are j-indexed, because of the different productivity levels under the two systems, which ultimately depend on the level of consumption. While in the market regime workers are type A agents, the whole population, including B types, forms the working class under the communist system ($c^A_{C,t} = c^B_{C,t} = c_{C,t}$). This implies that even if consumption is low in the market system, by construction it does not change over time (see below). Therefore, productivity $A_M(c^A_{M,t})$ remains constant, while it will become apparent below that $A_C(c_{C,t})$ will decrease over time. This will be, in fact, the main social learning outcome after experimenting with an unknown economic system abolishing private property and market signals: all agents will gradually discover that communism forces equality of ownership and centralized allocation that come at the cost of low efficiency and poor coordination.

The incentive structure under the two economic systems is captured in our model by the *(mis)alignment* of *ownership* and *control*. This is in line with the large literature on the key weaknesses of socialism: one strand dealing with ownership, control and economic incentives under 'soft budget constraints' (e.g., Kornai, 1980) – what Roemer (2009) labels 'lack of incentives'; another pointing to the overambitious task of central planning, given 'dispersed and local information', to ensure better allocative decisions than markets (e.g., von Hayek, 1940, 1945) – what Roemer (2009) labels 'lack of coordination'. Our approach highlights these two familiar disadvantages of a communist economy at their crucial link, the *intertemporal optimization decision*, at which the (mis)alignment of ownership and control manifests itself. The choice of consumption and accumulation out of one's own wealth given the signals of competitive markets and locally relevant information under capitalism sustains efficiency but generates inequality. 'Delegating' this choice to a 'principal' (be it the central planner or the communist party), suffering from an informational distance to production inputs and consumption needs, forces equality by revolution but erodes economic efficiency, thus making everyone equal in their poverty.

2.4 Income and Capital Accumulation

In the market system, income for capital owners and workers in each period is respectively

$$Y^B_{M,t} = r_{M,t} s_{t-1} = \alpha A_M(c^A_{M,t}) s_{t-1},$$

where s_{t-1} is savings in the previous period, and

$$Y^A_{M,t} = w_{M,t} = (1-\alpha) A_M(c^A_{M,t}).$$

We consider a satiation consumption level $\bar{c}$, under which savings do not occur. For ease of exposition, we impose a scenario where only the capital owners can invest, i.e.

$$\begin{aligned} Y^A_{M,t} &\leq \bar{c}, \\ Y^B_{M,t} &> \bar{c}. \end{aligned}$$

Capitalists choose savings to maximize utility in (2) given the budget constraint

$$c^B_t + s^B_t \leq Y^B_{M,t}. \tag{6}$$

To sum up, the timing of events during the accumulation process is as follows: yields from the savings of the previous period $r_M s_{t-1}$ comprise the capital stock of the present period K_t, which will then be put into production along with labor given the production function in (5); the output is then divided between consumption c_t^B and savings s_t^B (forming the future capital stock).

Under communism, capital is nationalized, i.e. capitalists are deprived of their ownership, and investment decisions are no longer individual but made by the central planner. As a consequence, individual income becomes a centralized allocation of an equal share of output to each member of the society, i.e. wage. In the aggregate, output still has to be equal to the sum of consumption and investment, their relative share being decided by the central planner, hence

$$C_t + S_t = H_{C,t}. \tag{7}$$

Therefore, under communism each individual gets the same consumption level, equal to

$$c_{C,t} = \frac{H_{C,t} - S_t}{1} = C_t.$$

The same timing holds for the accumulation process under the communist regime: $r_C S_{t-1}$ comprises K_t, which is used for national production $H_{C,t}$, the output of which is allocated between consumption in the society C_t and further savings S_t. Note that in this regime there is no market price of capital, therefore r_C is the shadow price of capital referred to by the central planner.

2.5 Productivity and Efficiency

The savings decision by a central planner differs from private savings decisions by individuals in that aggregate values are considered. Aggregate consumption is then divided equally among all agents. To compute the relative efficiency of the two regimes, $\chi_{C,t}$, we first note the budget constraint of an individual versus the one of the central planner:

$$\begin{aligned} c_t^B + s_t^B &\leq Y_{M,t}^B = \alpha A_M(c_{M,t}^A) s_{t-1}^B, \\ C_t + S_t &\leq H_{C,t} = A_C(c_{C,t})[\alpha K_t + (1-\alpha)L] = A_C(c_{C,t})[\alpha S_{t-1} + (1-\alpha)L]. \end{aligned} \tag{8}$$

It is easy to see from (8) that under a centralized system consumption must be cut for all if aggregate savings are to increase. This may differ from the optimal consumption decisions made by private individuals.

Our index of relative efficiency can be expressed in terms of potential growth possibilities under the two regimes. In the communist system, this can only be reached by increasing the total output, while in a market system it stems from the savings decision by capital owners. Hence, a comparison of the per capita values of the right-hand side

of the two budget constraints in (8) will deliver

$$\begin{aligned}\chi_{C,t} &= \frac{Y^B_{M,t}}{H_{C,t}} = \frac{c^B_t + s^B_t}{C_t + S_t} = \frac{\alpha A_M(c^A_{M,t})\hat{s}^B_{t-1}}{\alpha A_C(c_{C,t})S^*_{t-1} + (1-\alpha)A_C(c_{C,t})L} \\ &= \frac{A_M(c^A_{M,t})}{A_C(c_{C,t})}\frac{\hat{s}^B_{t-1}}{S^*_{t-1} + \frac{1-\alpha}{\alpha}L}, \end{aligned} \tag{9}$$

where $\hat{s}^B_{t-1}$ denotes a notional value computed by individuals should the market system be operative,[6] and S^*_{t-1} the optimal savings chosen by the central planner.

From equation (9), an increase in S^*_t leads to an acceleration of the accumulation process under communism, hence a fall in $\chi_{C,t}$, potentially implying an initial gain in efficiency with respect to the market system. However, individuals may not be satisfied with their consumption allocation as they suffer from an externality linked to the transfer of decisions to an upper level. Should the central planner assign a lower consumption level to all of them, work incentives and, therefore, productivity under the communist regime, $A_C(c_{C,t})$, will be reduced. If the planner then responds by further increasing savings, to compensate for the reduced productivity, he only exacerbates the relative inefficiency of the communist regime making it less and less sustainable. Hence, with abolished private property incentives and market prices, the system by itself tends to become unsustainable in the medium-to-long run. Essentially, such a set-up resembles the overinvestment experience in communist countries during their period of initial industrialization and subsequent attempts to increase future production (and, in historical context, catch up with the West).

2.6 Intergenerational Transmission of Beliefs

Agents live for two periods. During childhood (in the first period of life), they are 'socialized' and acquire a particular type just when becoming adult. When mature (in the second period of life), they perform active economic and ideological roles in the society, and die at the end of the period, bequesting any capital wealth they may have accumulated.

We assume that type A agents always teach a communist ideology to their offspring to abolish inequality, while type B agents always teach a pro-market ideology favoring efficiency. This is a first channel of transmitting beliefs that captures the influence on ideology intensity within the family, and corresponds to what is termed 'vertical transmission' in the literature (Bisin and Verdier, 2001, 2010). The evolution over time of the relative degree of ideological determination to change the status quo, however, is also affected outside the family. This second channel operates through the influence on ideology intensity by peers and the broader environment, and is known as 'oblique transmission'.

Furthermore, the dynamics of the probability of a regime shift are endogenous to the present economic situation and depend on the disutility brought about by differences

[6] This notional value can be derived, even if imperfectly, by observing the rest of the world where the franchise was extended to prevent revolution (Acemoglu and Robinson, 2000).

in income (our proxy for material well-being). What socialization efforts $\tau_t^i(\cdot)$ affect is the determination of the next generation to mobilize in order to change the system. Socialization effort by type A is generated by resentment from internal inequality υ_t:

$$\tau_t^A = \tau_t^A(\upsilon_t),$$

and, for type B, by the inferior efficiency with respect to the rest of the world χ_t:

$$\tau_t^B = \tau_t^B(\chi_t).$$

The properties of these socialization functions are

$$\tau_t^i(1) = 0,\ \tau^{i\,\prime}(\cdot) > 0,\ \tau^{i\,\prime\prime}(\cdot) < 0.$$

[Figure 2 about here]

Figure 2 presents the socialization process of the agents of each type A and B. The transition probabilities at time t, $P_t^{i\uparrow}$, that a parent of type i has a child with a *stronger* ($\uparrow$) or *weaker* ($\downarrow$) ideological determination can be written as

$$\begin{aligned} P_t^{A\uparrow} &= \tau_t^A(\upsilon_t) + [1 - \tau_t^A(\upsilon_t)]q_t; \\ P_t^{A\downarrow} &= [1 - \tau_t^A(\upsilon_t)]\,(1 - q_t)\,; \\ P_t^{B\uparrow} &= \tau_t^B(\chi_t) + [1 - \tau_t^B(\chi_t)]\,(1 - q_t)\,; \\ P_t^{B\downarrow} &= [1 - \tau_t^B(\chi_t)]q_t. \end{aligned} \tag{10}$$

Given these transition probabilities, the relative strength of individuals of type A in period $t+1$ is

$$\begin{aligned} q_{t+1} &= q_t P_t^{A\uparrow} + (1 - q_t)P_t^{B\downarrow} \\ &= q_t + (q_t - q_t^2)[\tau_t^A(\upsilon_t) - \tau_t^B(\chi_t)]. \end{aligned} \tag{11}$$

The properties of the socialization functions imply that $\tau_t^A(\upsilon_t) = 0$ under communism while $\tau_t^B(\chi_t) = 0$ under a market economy. Accordingly, the law of motion in (11) simplifies to

$$q_{M,t+1} = q_{M,t} + (q_{M,t} - q_{M,t}^2)\tau_t^A(\upsilon_{M,t}) \tag{12}$$

under markets and to

$$q_{C,t+1} = q_{C,t} - (q_{C,t} - q_{C,t}^2)\tau_t^B(\chi_{C,t}) \tag{13}$$

under communism.

Thus, in the market system the degree of ideological determination of type A (relative to type B) to change the status quo, $q_{M,t}$, increases with any positive socialization effort by type A, $\tau_t^A(\upsilon_{M,t}) > 0$. Above the critical value $q_{M,T} = 0.5$, type A's are

strong enough to overthrow the capitalist regime. In the communist system, initially aggregate savings S_t^* increases, hence relative output $\chi_{C,t}$ may even increase. However, since productivity depends positively on consumption, which in turn depends negatively on savings according to the aggregate budget constraint in (7), eventually the negative incentive effect through a lower $A_C(c_{C,t})$ dominates the positive effect of the increase in savings S_t^* (see equation (9)). This triggers reactions by type B agents, who increase their socialization effort $\tau_t^B(\chi_{C,t})$. In the communist system, therefore, the ideological determination of type B (relative to type A) $1-q_{C,t}$ to change the status quo increases, i.e. $q_{C,t}$ decreases, with any positive socialization effort by type B, $\tau_t^B(\chi_{C,t}) > 0$. As type B agents promote market values among type A agents (see (10)), who react by switching type as $q_{C,t}$ evolves over time, type B's succeed to bring down the communist regime once they reach the critical value $1 - q_{C,T+n} = 0.5$.

3 Dynamics: From Inequality to Inefficiency and Back

In the two economic systems there is a logical order in the *sequencing of actions*. In both systems the agent type who has the ownership and control (economic power) to decide on the split between consumption and saving/investment moves first, thus determining the economic allocations in equilibrium. The agent type who has no ownership and control rights and interests, thus no economic power, can only have socialization (or ideological) power by preaching against the regime in force, that is, trying to teach the next generation in favour of his/her own beliefs/values.

3.1 Market-Based Economy

In the market system, capitalists both own (legally) and control (effectively) the allocation of their income between consumption and savings to be bequested to their offspring, hence invested, and used to produce in the next period. In contrast, workers do not own and control anything apart from their labor force, which they supply inelastically in the model. It is therefore optimal saving and capital accumulation within the capitalist 'dynasties' (where ownership and control rights and interests are *aligned and effective*) that drives the efficiency and sustainability of the market system, but deepens the social inequality.

The type-B agents in this case are the first movers in a von Stackelberg leadership game and decide on savings, while taking into consideration in their maximization problem the socialization reaction of type-A agents to the inequality caused by their own savings. Starting with type-A agents (the working class), they take savings as fixed and maximize their utility, (1):

$$\max_{\tau_t^A(\upsilon_{M,t})} U_{M,t}^A(\cdot) = c_{M,t}^A + \beta Y_{M,t+1}^A - E_t^A(\upsilon_{t+1}) - \frac{\tau_t^A(\upsilon_{M,t})^2}{2}.$$

After a few substitutions (see Appendix B.1), we can rewrite the optimization prob-

lem as:

$$\max_{\tau_t^A(\upsilon_{M,t})} (1+\beta)(1-\alpha) A_M(c_{M,t}^A) \\ -\left\{1-\left[q_t + q_t(1-q_t)\tau_t^A(\upsilon_t)\right]\right\} \ln\left(\tfrac{\alpha}{1-\alpha} s_t^B\right) - \tfrac{\tau_t^A(\upsilon_{M,t})^2}{2}.$$

The first-order condition yields the optimal reaction of type A as follower:

$$\begin{aligned} \frac{\partial U_{M,t}^A(\cdot)}{\partial \tau_t^A(\upsilon_{M,t})} &= q_t(1-q_t)\tau_t^{A\,\prime}(\upsilon_{M,t}) \ln\left(\frac{\alpha}{1-\alpha} s_t^B\right) - \tau_t^A(\upsilon_{M,t})\tau_t^{A\,\prime}(\upsilon_{M,t}) = 0 \\ &\Leftrightarrow \quad \tau_t^A(\upsilon_{M,t})^* = q_t(1-q_t) \ln\left(\frac{\alpha}{1-\alpha} s_t^B\right). \end{aligned} \tag{14}$$

This equation delivers a preliminary insight on the mechanisms that drive the evolution from one system to another. It is easy to see that an increase in the private savings (by the capitalists) leads to increased socialization effort by type A agents. The latter can only expect a growing inequality between the two types of agents, which reinforces their determination to change the regime. More precisely, the higher the expected inequality, the higher the effort to transmit their preferences towards a more equal society, i.e. communism.

Turning to type B agents, they move first by making a decision on the amount of their savings:

$$\begin{aligned} \max_{s_{M,t}^B} U_{M,t}^B(\cdot) &= c_{M,t}^B + \beta Y_{M,t+1}^B - E_t^B(\chi_{t+1}) - \frac{\tau_t^B(\chi_{M,t})^2}{2} \\ \text{s.t. } c_{M,t}^B + s_{M,t}^B &\leq Y_{M,t}^B. \end{aligned}$$

We make substitutions again (see Appendix B.1) and rewrite, omitting below the M-subscript to savings because under communism individual savings are absent:

$$\max_{s_t^B} \left(Y_{M,t}^B - s_t^B\right) + \beta r_M s_t^B - q_{t+1} \ln(\chi_{C,t+1}).$$

Notice that we take $\chi_{C,t+1}$ as given. This implies that capital owners in a free market system use a notional savings rate when considering their expected loss in the next period from inefficiency brought about by the alternative system.[7]

After further substitutions (see Appendix B.1), we obtain:

$$\max_{s_t^B} \quad \alpha A_M(c_{M,t}^A) s_{t-1}^B + \left(\beta\alpha A_M(c_{M,t}^A) - 1\right) s_t^B - \left[q_t + (q_t - q_t^2)\tau_t^A(\upsilon_t)\right] \ln\left(\chi_{C,t+1}\right).$$

[7] With an endogenous $\chi_{C,t+1}$, higher savings would increase the disutility of a regime change to capital owners since current savings would appear in the numerator of $\chi_{C,t+1}$ creating a larger gap between output under the two systems.

Replacing for $\tau_t^A(\upsilon_t)$ with the optimal reaction of type A agents from (14) delivers

$$\begin{aligned}&\max_{s_t^B} \quad \alpha A_M(c_{M,t}^A)s_{t-1}^B + \left(\beta\alpha A_M(c_{M,t}^A) - 1\right) s_t^B \\ &- \left[q_t + q_t^2(1-q_t)^2 \ln\left(\frac{\alpha}{1-\alpha}s_t^B\right)\right] \ln\left(\chi_{C,t+1}\right).\end{aligned}$$

The first-order condition yields optimal savings by type B as leader:

$$\begin{aligned}\frac{\partial U_{M,t}^B(\cdot)}{\partial s_t^B} &= \beta\alpha A_M(c_{M,t}^A) - 1 - \frac{q_t^2(1-q_t)^2}{s} \ln\left(\chi_{C,t+1}\right) = 0 \\ &\Leftrightarrow \quad s_t^{B*} = q_t^2(1-q_t)^2 \frac{\ln\left(\chi_{C,t+1}\right)}{\beta\alpha A_M(c_{M,t}^A) - 1}, \qquad (15)\end{aligned}$$

where $A_M(c_{M,t}^A) > 1/\beta\alpha$ must hold for positive savings by the capital owners.

The last equation reveals that increased expected inefficiency under the alternative (communist) system induces higher accumulation by capital owners. In addition, the higher the productivity (A_M), the lower the need to save.

Substituting (15) back into (14) to derive the optimal socialization effort of the type A in its final form, we get

$$\tau_t^A(\upsilon_{M,t})^* = q_t(1-q_t)\ln\left(\frac{\alpha q_t^2(1-q_t)^2 \ln\left(\chi_{C,t+1}\right)}{(1-\alpha)[\beta\alpha A_M(c_{M,t}^A) - 1]}\right). \qquad (16)$$

Substituting $\tau_t^A(\upsilon_{M,t})^*$ from (16) into (11), next-period ideological determination of type A to change the status quo becomes

$$\begin{aligned}q_{t+1} &= q_t + (q_t - q_t^2)\tau_t^A(\upsilon_{M,t})^* \\ &= q_t + q_t^2(1-q_t)^2 \ln\left(\frac{\alpha q_t^2(1-q_t)^2 \ln\left(\chi_{C,t+1}\right)}{(1-\alpha)[\beta\alpha A_M(c_{M,t}^A) - 1]}\right). \qquad (17)\end{aligned}$$

It is seen from (17) that the evolution of q_t with time under a market system takes a positive value when $\tau_t^A(\upsilon_{M,t})^* > 0$. This is true as long as

$$\alpha q_t^2(1-q_t)^2 \ln\left(\chi_{C,t+1}\right) > (1-\alpha)[\beta\alpha A_M(c_{M,t}^A) - 1].$$

In sum, the capital owners allocate their income between consumption and savings, the type A agents then react by choosing their socialization effort, which in turn affects the ideological determination and strength of the next generation of A types to change the status quo.

We highlight our key analytical results underlying the transition from a market-based to a centrally-planned economic system in the following proposition:

Proposition 1 (Communist Revolution) *Assume first that a change of system in our model occurs once the strength of the follower type (A) dominates that of the leader type (B), i.e.* $q_t > 0.5$. *Then, for any lower initial degree of strength of type A relative to type B,* $q_0 < 0.5$ *under which a market regime is feasible, given their optimal socialization effort,* $\tau_t^A(\upsilon_{M,t})^*$, *and the law of motion of* q_t, *the optimal savings by type B's,* s_t^{B*}, *are* increasing *in* q_t. *Such an equilibrium strategic behavior in the repeated von Stackelberg game we consider generates a communist revolution in some period* T *as soon as* $q_T > 0.5$.

Proof of Proposition 1. We are here interested to sign the effect of an increase in the strength of type A's ($q_{t+1} > q_t$) on the optimal savings by capital owners. This effect is given by the respective partial derivative:

$$\frac{\partial s_t^{B*}}{\partial q_t} = 2q_t(1-q_t)(1-2q_t)\frac{\ln\left(\chi_{C,t+1}\right)}{\beta\alpha A_M(c_{M,t}^A)-1} \begin{array}{ll} >0 & \text{if } q_t < 0.5 \\ <0 & \text{if } q_t > 0.5 \end{array}. \tag{18}$$

Thus, for any low $q_t < 0.5$, $\frac{\partial s_t^{B*}}{\partial q_t} > 0$. In words, the optimal reaction functions of the two types in the repeated von Stackelberg game of class struggle under markets lead to a progressive increase of q_t until it surpasses some 'critical mass' or threshold beyond which type A's become sufficiently strong and ideologically determined to overthrow the existing capitalist social order by communist revolution. This happens as soon as $q_T > 0.5$. ■

Our interpretation of Proposition 1 is the following: a market system is only feasible for $q_0 < 0.5$. An increase in this strength caused by social resentment induces more savings by capitalists as they perceive a higher threat of losing power. This increases the efficiency of the system, thus the profits, augmenting the revenues of the capital owners; but it also increases inequality, feeding the resentment of type A agents, and ultimately increasing the probability of a regime change. While the workers gradually become stronger and more ideologically motivated to overthrow the system, capitalists accumulate more. Increased savings lead to the uprising of the working class. Hence, the market system moves towards its fall as resentment within type A agents has a reinforcing effect on capital accumulation, and inequality, rather than mitigating it. Historically, this seems to have been the case of Russia, and then the expansion of communism to Eastern Europe, China and elsewhere, on which we focus here.

Alternatively, if we relax the assumption of a regime change at $q_t = 0.5$, once the strength of type A's passes a threshold beyond which the probability of a regime change is perceived by type B's as credible ($q_t > 0.5$), the latter would adapt their behavior.

Corollary 1 (Stabilizing Capitalism) *Assume alternatively that an increase in* q_t *beyond the threshold* $q_t = 0.5$ *may not necessarily generate a communist revolution. Then, for any degree of strength of type A's relative to type B's higher than* 0.5 *under markets, given their optimal socialization effort,* $\tau_t^A(\upsilon_{M,t})^*$, *and the law of motion of* q_t, *the optimal savings by type B's,* s_t^{B*}, *are* decreasing *in* q_t. *The equilibrium strategic*

behavior in the repeated von Stackelberg game we consider now reduces the accumulation of capital and, in effect, the degree of inequality, ultimately stabilizing the capitalist market-economy system.

Proof of Corollary 1. Follows from (18) above. ■

Corollary 1 describes certain more intricate dynamics under a market economy: for relatively strong type A's, $0.5 < q_t < 1$, given the optimal socialization effort of type A and the law of motion of q_t, an increase in the probability of a regime change ($q_{t+1} > q_t$) induces a reduction in savings by capital owners, which reduces efficiency, profits and, hence, income inequality. By reducing the ideological determination of type A's to change the status quo, such a reaction by capital owners can avoid slipping towards communism. Historically, this seems to have been the case of the Western world and the welfare state, where democratization of capitalism and redistribution of income have preserved the market system (although modified). As we noted, this scenario is not at the center of our interest, and has been analyzed elsewhere in the literature.

3.2 Centrally-Planned Economy

Under communism, no one makes economic decisions apart from the central planner, who is of type A and splits total consumption equally across all members of society.[8] The latter are nominally or notionally (de jure) owners of all the capital stock after the nationalization following the communist revolution and the control is delegated to the communist party or central planner. However in practice (de facto) individuals do not control the choice of aggregate consumption and investment out of national income, which is also national output. Thus, under communism, there is *misalignment* of ownership and control rights and interests between the individuals, equal in their consumption allocation, and the central planner or the communist leader, who makes the economic choices in the name of the equal society. If inconsistent with what individuals would have chosen to do by themselves, this creates inefficiency. We capture and interpret it in comparing the optimization under central planning (aggregate, then disaggregated top-down by equal split) vis-à-vis the market (individual, aggregated bottom-up).

At the moment of transition, T, the total amount of savings by capital owners in the last period of the market system $(1 - q_{S,0})s^{B,*}_{T-1}$ defines the available capital stock under communism – see Appendix A.1. In subsequent periods, the central planner will maximize its utility in the name of the type A agents taking into account aggregate values. Therefore, the central planner (not individual capitalists, whose capital has been nationalized) optimally chooses the level of aggregate savings, i.e. national investment. This also determines the allocation of output to be distributed equally among the total

[8] Note that we ignore neither that inequalities were de facto existing in communist countries, nor that they were creating resentment (see, e.g., Joo, 2005, for an account). However, considering explicitly the nomenklatura would only complicate the model without changing the substance of the results (in effect, only accelerating the swing back from plan to market).

population for consumption.[9]

The central planner here is the first mover and decides on savings. In this aggregate optimization problem, he takes into consideration the socialization reaction of type B agents to the inefficiency caused by the centrally chosen savings. Starting with type B agents, they take savings as fixed and maximize:

$$\begin{aligned}\max_{\tau_t^B(\chi_{C,t})} U_{C,t}^B(c_{C,t}^B, Y_{C,t+1}^B, \chi_{t+1}) &= c_{C,t}^B + \beta Y_{C,t+1}^B - E_t^B(\chi_{t+1}) - \frac{\tau_t^B(\chi_{C,t})^2}{2},\\ \max_{\tau_t^B(\chi_{C,t})} U_{C,t}^B(c_{C,t}^B, Y_{C,t+1}^B, \chi_{t+1}) &= -E_t^B(\chi_{t+1}) - \frac{\tau_t^B(\chi_{C,t})^2}{2}.\end{aligned}$$

After analogous substitutions (see Appendix B.2), we obtain:

$$\max_{\tau_t^B(\chi_{C,t})} - q_{t+1} \ln \chi_{C,t+1} - \frac{\tau_t^B(\chi_{C,t})^2}{2},$$

$$\max_{\tau_t^B(\chi_{C,t})} - \left\{q_t + q_t(1-q_t)[-\tau_t^B(\chi_t)]\right\} \ln\left(\frac{\alpha A_M(c_{M,t}^A) s_t^B}{\alpha A_C(c_{C,t}) S_t}\right) - \frac{\tau_t^B(\chi_{C,t})^2}{2}.$$

The first-order condition then yields:

$$\begin{aligned}\frac{\partial U_{C,t}^B(\cdot)}{\partial \tau_t^B(\chi_{C,t})} &= q_t(1-q_t)\tau_t^{B\,\prime}(\chi_{C,t}) \ln\left(\frac{A_M(c_{M,t}^A) s_t^B}{A_C(c_{C,t}) S_t}\right) - \tau_t^B(\chi_{C,t})\tau_t^{B\,\prime}(\chi_{C,t}) = 0\\ &\Leftrightarrow \quad \tau_t^B(\chi_{C,t})^* = q_t(1-q_t) \ln\left(\frac{A_M(c_{M,t}^A) s_t^B}{A_C(c_{C,t}) S_t}\right). \qquad (19)\end{aligned}$$

Hence, an increase in the planner's aggregate savings reduces the socialization effort by type-B agents. This is due to the fact that, as seen above, such an increase may give a boost, at least initially, to the productivity of the communist system. While this effect per se could reduce inefficiency, this accumulation process has the adverse effect of reducing consumption opportunities available to all workers. This decreases incentives $A_C(c_{C,t})$ and productivity with it. As a consequence, the total effect of aggregate savings on productivity and socialization effort by agents who share the values of capital owners is ambiguous. However, once the negative incentive effect outweighs the direct positive impact of savings on productivity, inefficiency increases and type-B agents *recruit* intertemporally by increasing their socialization effort.

The central planner as a first mover maximizes utility in the name of the type A agents taking into account aggregate values. Therefore, the central planner (not individual capitalists, whose capital has been nationalized) optimally chooses the level of aggregate savings, i.e. national investment. This also determines the allocation of output to be distributed equally among the total population for consumption.

[9] This follows our assumption of inequality-aversion among type A agents, to conform with the preference for equality among the thinkers and pioneers of communism.

$$\begin{aligned}\max_{S_t} \quad U^A_{C,t}(C_t, K_{C,t+1}, \upsilon_{t+1}) &= C_t + \beta K_{C,t+1} - E^A_t(\upsilon_{t+1}) - \frac{\tau^A_t(\upsilon_{C,t})^2}{2} \\ \text{s.t. } C_t + S_t &\leq H_{C,t}.\end{aligned}$$

Note that, parallel to the market economy, where $Y_{M,t+1}$ was the basis for private investment decisions, the central planner uses $K_{C,t+1} = \alpha A_C(c_{C,t})S_t$ in his optimization problem. Substituting further (see Appendix B.2), we obtain:

$$\begin{aligned}&\max_{S_t} \quad A_C(c_{C,t})[\alpha S_{t-1} + (1-\alpha)L] - (1 - \beta\alpha A_C(c_{C,t}))S_t \\ &- \left(1 - \left[q_t - (q_t - q_t^2)\tau^B_t(\chi_t)\right]\right)\ln\left(\upsilon_{M,t+1}\right).\end{aligned}$$

Replacing for $\tau^B_t(\chi_t)$ with the optimal reaction curve by type B agents derived in (19) we get

$$\begin{aligned}&\max_{S_t} \quad A_C(c_{C,t})[\alpha S_{t-1} + (1-\alpha)L] - (1 - \beta\alpha A_C(c_{C,t}))S_t \\ &- \left(1 - \left[q_t - q_t^2(1-q_t)^2 \ln\left(\frac{A_M(c^A_{M,t})s^B_t}{A_C(c_{C,t})S_t}\right)\right]\right)\ln\left(\upsilon_{M,t+1}\right).\end{aligned}$$

The first-order condition yields:

$$\begin{aligned}\frac{\partial U^A_{C,t}(\cdot)}{\partial S_t} &= (\beta\alpha A_C(c_{C,t}) - 1) + \frac{q_t^2(1-q_t)^2\ln\left(\upsilon_{M,t+1}\right)}{S_t} = 0 \\ &\Leftrightarrow \quad S^*_t = \frac{q_t^2(1-q_t)^2\ln\left(\upsilon_{M,t+1}\right)}{1 - \beta\alpha A_C(c_{C,t})},\end{aligned} \tag{20}$$

where $A_C(c^A_{M,t}) < \frac{1}{\beta\alpha}$ must hold for positive savings by the the central planner.

The last equation notably reveals that increased expected inequality under the alternative (market) system induces higher savings by the central planner in an effort to further root the communist system. In addition, the higher the productivity (A_C), the lower the need to save.

Substituting (20) back into (19) to derive the optimal socialization effort of type B in its final form, we get

$$\tau^B_t(\chi_{C,t})^* = q_t(1-q_t)\ln\left(\frac{A_M(c^A_{M,t})s^B_t[1 - \beta\alpha A_C(c_{C,t})]}{A_C(c_{C,t})q_t^2(1-q_t)^2\ln\left(\upsilon_{M,t+1}\right)}\right). \tag{21}$$

And now substituting $\tau^B_t(\chi_{C,t})^*$ from (21) into (11), next-period proportion of the population in favor of a market-based system becomes

$$\begin{aligned}q_{t+1} &= q_t - (q_t - q_t^2)\tau^B_t(\chi_{C,t})^* \\ &= q_t - q_t^2(1-q_t)^2\ln\left(\frac{A_M(c^A_{M,t})s^B_t[1 - \beta\alpha A_C(c_{C,t})]}{A_C(c_{C,t})q_t^2(1-q_t)^2\ln\left(\upsilon_{M,t+1}\right)}\right).\end{aligned} \tag{22}$$

It immediately appears from (22) that the ideological stance of type A's relative to type B's under the communist system weakens when $\tau_t^B(\chi_{C,t})^* > 0$. This is true as long as

$$A_M(c_{M,t}^A)s_t^B[1-\beta\alpha A_C(c_{C,t})] > A_C(c_{C,t})q_t^2(1-q_t)^2\ln(\upsilon_{M,t+1}).$$

As discussed earlier, this depends on the total effect of aggregate savings by the central planner on productivity. Indeed, although savings directly increase productivity in support of communism, expression

$$\frac{\partial q_{t+1}}{\partial A_C(c_{C,t})}=\frac{q_t^2(1-q_t)^2}{A_C(c_{C,t})[1-\beta\alpha A_C(c_{C,t})]}>0 \tag{23}$$

shows that the disincentives created by reduced consumption due to more government savings increases the relative strength of type B's to change the status quo (now also *proportion* of the population who sympathize with communism, as more agents convert to the type-B ideology).[10]

In sum, the central planner allocates national income between consumption and savings at the aggregate level, the type B agents then react by choosing their socialization effort to influence the ideology of the rest of the society (type A). The proportion of type B agents in the next period $(1-q_{t+1})$ is then determined.

We highlight our key analytical results underlying the transition from a centrally-planned to a market-based economic system in another proposition:

Proposition 2 (Market Transition) *Assume again that a change of system in our model occurs once the strength of the follower type (B) dominates that of the leader type (A), i.e.* $q_t < 0.5$. *Then, for any lower initial degree of strength of type B relative to type A,* $q_T > 0.5$ *above which a communist regime is feasible, given the optimal socialization effort by type B,* $\tau_t^B(\chi_{C,t})^*$, *and the law of motion of* q_t, *the optimal savings by the central planner (in the name of type A's),* S_t^*, *are* decreasing *in* q_t *(that is, increasing in* $1-q_t$*). Such an equilibrium strategic behavior in the repeated von Stackelberg game we considered generates transition to a market-based economy in some period* $T+n$ *as soon as* $q_{T+n} < 0.5$.

Proof of Proposition 2. We now have to sign the effect of a decrease in the strength of type A's relative to type B's ($q_{t+1} < q_t$) on aggregate savings by the central planner. This effect is given by the respective partial derivative:

$$\frac{\partial S_t^*}{\partial q_t}=2q_t(1-q_t)(1-2q_t)\frac{\ln(\upsilon_{M,t+1})}{1-\beta\alpha A_C(c_{C,t})}\begin{array}{ll}>0 & \text{if } q_t<0.5\\ <0 & \text{if } q_t>0.5\end{array}. \tag{24}$$

Thus, for any high $q_t > 0.5$ (i.e. any low $1-q_t < 0.5$), $\frac{\partial S_t^*}{\partial q_t} < 0$ (i.e. $\frac{\partial S_t^*}{\partial(1-q_t)} > 0$). In words, the optimal reaction functions of the two types in the repeated von Stackelberg

[10] Note that $\frac{\partial \tau_t^B(\chi_{C,t})^*}{\partial A_C(c_{C,t})} < 0$, that is socialization effort by type B increases due to decreased consumption, incentives, and productivity.

game of conflicting beliefs under communism lead to a progressive increase of $1-q_t$, i.e. a progressive decrease of q_t until it drops below certain 'critical mass' or threshold beyond which type B's become sufficiently strong and ideologically determined to overthrow the existing communist social order by pro-market transition reforms. This happens as soon as $q_{T+n} < 0.5$. ■

Our interpretation of Proposition 2 is the following: a communist system is only feasible for $1 - q_T < 0.5$. An increase in this strength caused by a shift of ideology that arises from lower relative productivity under communism induces more savings by the central planner. He, thus, responds by attempting to overcome through investment and accumulation the undesirable inefficiency of the communist system gradually being discovered in the wake of the revolution with nationalization as communism operates period after period. Yet, the decision being made at an aggregate level, it reduces total available consumption equally divided within the society. Lower consumption (material well-being) leads to lower incentives by workers, and hence an adverse effect on productivity. As this ultimately increases inefficiency, type B agents respond by higher socialization effort, and more discontented people who have now observed the consequences of communism shift ideology to support a market system. This trend continues until the point where economic transition is triggered, $q_{T+n} < 0.5$, and the regime reverts to the market system. Historically, this seems to have been the Soviet and East European case, on which we focus here.

Again, if we drop the assumption that a regime change occurs at threshold $q_t = 0.5$, once the strength of type B's drops to a threshold beyond which the probability of a regime change is perceived by type A's as credible ($q_t < 0.5$), the latter would adapt their behavior.

Corollary 2 (Stabilizing Communism) *Assume alternatively that a reduction in q_t below the threshold $q_t = 0.5$ may not necessarily generate a pro-market transition. Then, for any degree of strength of type B relative to type A higher than* 0.5 *under communism, given their optimal socialization effort, $\tau_t^B(\chi_{C,t})^*$, and the law of motion of q_t, the optimal savings by the central planner (in the name of type A's), S_t^*, are* increasing *in q_t (that is, decreasing in $1 - q_t$). The equilibrium strategic behavior in the repeated von Stackelberg game we consider now reduces the accumulation of capital and, in effect, the degree of inefficiency, ultimately stabilizing the communist centrally-planned economic system.*

Proof of Corollary 2. Follows from (24) above. ■

Corollary 2 describes certain more intricate dynamics under a communist economy: for a relatively high fraction of type B's, $0.5 < 1-q_t < 1$, given the optimal socialization effort of type B and the law of motion of q_t, an increase in the probability of a regime change ($q_{t+1} < q_t$) induces a reduction in aggregate savings by the central planner, which increases consumption allocations (material well-being), hence productivity and,

ultimately, the efficiency of the communist system. By reducing the ideological determination of type B's to change the status quo, such a reaction by the central planner can prolong the life of a communist regime and, potentially, avoid surrendering central planning. Historically, this seems to have been the Chinese case, yet pro-market reforms were undertaken there widely in coexistence with the mono-party system.

4 Social Learning and Economic Change

Our theoretical analysis above summarized the rise and fall of communism as a process of social learning from experimenting with a new economic system that failed. It also showed how the same general mechanism we emphasized as driving social evolution could generate, under certain conditions and under minor regime-dependent specificity, not just the advent of communism but also its demise. That is, we proposed a model of long-run economic dynamics as one possible explanation for the principal insight from the history and the turn of events during the last century and a half.

Our work is in line with the main ideas in the recent book by North (2005) on the process of long-run economic change. In it North extends his earlier contributions to the new institutional economics. The need for such a quite substantial extension was obvious, as the neoclassical paradigm and its formal representation in general equilibrium theory were not intended, hence designed, 'to explain the process of economic change' (in the words of North, 2005, p. vii). He characterizes succinctly the nature of this social evolutionary process as follows:

> "In contrast to Darwinian evolutionary theory, the key to human evolutionary change is the intentionality of the players. The selection mechanisms in Darwinian evolutionary theory are not informed by beliefs about the eventual consequences. In contrast, human evolution is guided by the perceptions of the players; choices – decisions – are made in the light of those perceptions with the intent of producing outcomes downstream that will reduce uncertainty of the organizations – political, economic, social – in pursuit of their goals. Economic change, therefore, is for the most part a deliberate process shaped by the perceptions of the actors about the consequences of their actions. The perceptions come from the beliefs of the players – the theories they have about the consequences of their actions – beliefs that are typically blended with their preferences." (North, 2005, p. viii)

Among the other examples confirming the above view, North (2005) stresses in particular the experience with communism in Russia, which forms chapter 11 (pp. 146–154) in his book. In short, the "story of the Soviet Union is a story of perceived reality → beliefs → institutions → policies → altered perceived reality and on and on." (North, 2005, p. 4). While the book by North (2005) is framed along purely descriptive argumentation, in the tradition of the new institutional economics, the chain of logic in the last quote certainly reminds as well of another strand of quite technical, statistical literature, featuring learning (that could be Bayesian, social or of other methodological

strand or aggregation level). In essence, it was, in part, our aim when developing the model of the rise and fall of communism to formalize analytically the core of the chain of logic above. We did not wish to go into excessive details and sophistication at this first attempt toward such a broad task. Our guiding principle was, therefore, to keep the framework fairly general and the learning process by all agents in the experiment with communism under aggregate uncertainty – or, more precisely, ambiguity hidden within an unattempted ever economic system – quite straightforward. Of course, many more additional ingredients, considerations and complications could be built into the set-up presented. We ourselves believe there are a number of interesting and relevant avenues to enrich the basic model we developed. Yet our goal with this paper was to capture the 'perceived reality → beliefs → institutions → policies → altered perceived reality' chain North (2005) emphasized in words into a coherent and general theoretical construct capable to summarize the experiment of communism as social learning in the face of ambiguity highlighting the trade-off between equality and efficiency.

Indeed, our model begins with a 'perceived reality' that is unjust for our type A agents, as they are born unequal and poorer. Their 'beliefs' are thus shaped out by the ideal of achieving equality, and are propagated by socialization and the spread of ideology across society in our model. At this initial point in our model, however, the world has never operated a communist economic system, to which the A types strive. In other words, the society faces huge (aggregate) ambiguity if it decides to attempt a change in the status quo. The experimentation with communism can, in this light, be seen as the 'necessary evil' to pass through in order to learn (more) about (the properties of) an unknown form of socioeconomic organization. The experiment accordingly creates its own 'institutions' and 'policies', forcing equality in incomes and a central planning system to replace the role of capitalists and markets. But after repeating a few generation-spans of production and consumption, the social realities imposed by the revolution and nationalization turn out simply not to work. All members of the communist region suffer lower and lower material well-being due to misaligned incentives resulting from a distorted ownership and control structure. As people in the communist region get poorer and poorer, while observing as a reference point the rest of the world that has remained market-based (although, possibly, under considerable democratization and redistribution to mitigate the inequality problem of early capitalism) and is performing better, a drive to pro-market reforms – the 'altered perceived reality' – reverts the society back to sustainability. Although we stop modeling the chain of social evolution at this point, it certainly does not end here, but continues by experimenting and discovering 'on and on'.

5 Conclusion: Up the Spiral of Knowledge

Sometimes – if not often – in history, the society faces the unavoidable challenge to experiment with its own existence and future under huge ambiguity. With heterogenous agents, information sets, expectations and interests, it is not always easy to converge to a commonly shared plan, or at least hope for such a plan to possibly end up success-

fully. Doubts, conflicts and ideologies emerge naturally, values and institutions evolve, responding to evolving realities, experiences, learning. Indeed,

> "Men make their own history, but they do not make it just as they please: they do not make it under circumstances chosen by themselves, but under circumstances directly encountered, given and transmitted from the past. The tradition of all the dead generations weighs like a nightmare on the brain of the living."[11]

At times, the experiment discovers a positive outcome. And then society finds and settles into a new (again, temporary) equilibrium, until the next unprecedented vital change of the environment. However, when the outcome of such a social experiment is negative, the pendulum of history swings back. Yet this is not, in fact, back, as social evolution progresses along a spiral, whose circles constitute a gradation of hard-to-acquire knowledge.

References

[1] Acemoglu, Daron and James A. Robinson (2000), "Why Did the West Extend the Franchise? Democracy, Inequality, and Growth in Historical Perspective," *Quarterly Journal of Economics*, 115, 1167–1199.

[2] Bisin, Antonio and Thierry Verdier (2001), "The Economics of Cultural Transmission and the Dynamics of Preferences," *Journal of Economic Theory* 97, 298–319.

[3] Bisin, Antonio and Thierry Verdier (2010), "The Economics of Cultural Transmission and Socialization," in J. Benhabib, A. Bisin, and M. Jackson, eds., *Handbook of Social Economics*, Elsevier.

[4] Brown, Archie (2010), *The Rise and Fall of Communism*, Harper Collins.

[5] Kornai, János (1980), *Economics of Shortage*, Amsterdam: North-Holland.

[6] von Hayek, Friedrich A. (1940), "Socialist Calculation: The Competitive 'Solution'," *Economica* 7, 125–149.

[7] von Hayek, Friedrich A. (1945), "The Use of Knowledge in Society," *American Economic Review* 30, 519–530.

[8] Joo, Hyung-min (2005), "Narratives of Inequality Under Communism: Voices from Below," *Problems of Post-Communism* 52, 46–58.

[9] Lange, Oscar (1956 [1936]), "On the Economic Theory of Socialism," in B. Lippincott, ed., *On the Economic Theory of Socialism*, Minneapolis: University of Minnesota Press.

[11] Karl Marx, 1852, *The Eighteenth Brumaire of Louis Bonaparte*, Chap. 1.

[10] North, Douglass C. (2005), *Understanding the Process of Economic Change*, Princeton and Oxford: Princeton University Press.

[11] Roemer, John E. (1985), "Rationalizing Revolutionary Ideology," *Econometrica* 53, 85–108.

[12] Roemer, John E. (2009), "Socialism", in S. Durlauf and L. Blume, eds., *The New Palgrave Dictionary of Economics* (2nd ed.), Palgrave Macmillan.

[13] Saint-Paul, Gilles (2010), "Endogenous Indoctrination: Occupational Choices, the Evolution of Beliefs and the Political Economy of Reforms," *Economic Journal* 120, 325–353.

[14] Tinbergen, Jan (1960), "Optimum Savings and Utility Maximization over Time," *Econometrica* 28, 481–489.

A Ownership Shifts in Economic Transitions

A.1 Nationalization of Capital after the Communist Revolution

Revolution coupled with nationalization takes place at date T. The capital stock, K_T, the central planner now has to manage is inherited from the expropriated capital owners. Assuming that capital owners represent a given share, $1 - q_{S,0}$, of the adult population, it follows that

$$K_T \equiv (1 - q_{S,0}) s^{B,*}_{T-1}. \tag{A.1}$$

Hence, from the very first period of communism, the production function is written as

$$H_{C,t}(A_C, K_t, L) = A_C(c^A_{C,t})[\alpha K_t + (1 - \alpha)L]$$

for $t = T, ..., T + n$.

Consequently, at T, the relative efficiency of the two systems, χ_t, is given by

$$\begin{aligned} \chi_{C,T} &= \frac{Y^B_{M,T}}{H_{C,T}} = \frac{c^B_T + s^B_T}{C_T + S_T} = \frac{\alpha A_M(c^A_{M,t})\hat{s}^B_{T-1}}{\alpha A_C(c_{C,t})S^*_{T-1} + (1 - \alpha)A_C(c_{C,T})L} \\ &= \frac{A_M(c^A_{M,T})}{A_C(c_{C,T})} \frac{\hat{s}^B_{T-1}}{S^*_{T-1} + \frac{1-\alpha}{\alpha}L}, \end{aligned} \tag{A.2}$$

becoming afterwards, for all $t > T$,

$$\chi_{C,t} = \frac{A_M(c^A_{M,t})}{A_C(c_{C,t})} \frac{\hat{s}^B_{t-1}}{S^*_{t-1} + \frac{1-\alpha}{\alpha}L},$$

which is equation (9) in the main text. A hat indicates that agents have to refer to a notional value: the one that would be observed should the alternative system be operative – or, in fact, observed in the non-communist rest of the world.

A.2 Privatization of Capital in the Course of Pro-Market Reforms

Transition coupled with privatization takes place at date $T+n$. The capital stock, K_{T+n}, the new capital owners now have to manage is inherited from the recovery of ownership rights. Historically, several ways of ownership transfer have been tried in post-communist transition economies, from mass privatization to (former) nomenklatura grabbing. In our context, it suffices to assume that the new capital owners will be the agents most eager to get into this new role, i.e. type B agents. At $T+n$, they represent a proportion $1 - q_{S,T+n}$ of the adult population. It thus follows that

$$K_{T+n} \equiv (1 - q_{S,T+n-1})\, S^*_{T+n-1}, \tag{A.3}$$

ensuring that, from the very first period of the return to a market system, the production function is written as

$$H_{M,t}(A_M, K_t, L) = A_M(c^A_{M,t})[\alpha K_t + (1-\alpha)L]$$

for $t = T+n, T+n+1, ...$

B Derivation of the Model Solutions

B.1 The von Stackelberg Game under Market Capitalism

The type-B agents in this case are the first movers in a von Stackelberg leadership game and decide on savings. They take into consideration the socialization reaction of type-A agents to the inequality caused by their savings in their maximization problem. Starting with type-A agents (the working class), they take savings as fixed and maximize:

$$\max_{\tau^A_t(\upsilon_{M,t})} U^A_{M,t}(\cdot) = c^A_{M,t} + \beta Y^A_{M,t+1} - E^A_t(\upsilon_{t+1}) - \frac{\tau^A_t(\upsilon_{M,t})^2}{2}.$$

We first substitute: $c^A_{M,t} = Y^A_{M,t} = w_{M,t} = (1-\alpha)\,A_M(c^A_{M,t})$; $Y^A_{M,t+1} = Y^A_{M,t} = const$ because $s^A_t = 0$; the expected regime-dependent inequality $E^A_t(\upsilon_{t+1})$ from its definition after noting that $\ln \upsilon_{C,t+1} = 0$; and $\upsilon_{M,t+1}$ (from its definition).

$$\max_{\tau^A_t(\upsilon_{M,t})} Y^A_{M,t} + \beta Y^A_{M,t} - (1-q_{t+1}) \ln\left(\frac{Y^B_{M,t+1}}{Y^A_{M,t+1}}\right) - \frac{\tau^A_t(\upsilon_{M,t})^2}{2},$$

$$\max_{\tau^A_t(\upsilon_{M,t})} (1+\beta)(1-\alpha)\,A_M(c^A_{M,t}) - (1-q_{t+1}) \ln\left(\frac{Y^B_{M,t+1}}{Y^A_{M,t+1}}\right) - \frac{\tau^A_t(\upsilon_{M,t})^2}{2}.$$

Using $Y^B_{M,t+1} = \alpha A_M(c^A_{M,t}) s^B_t$ and $Y^A_{M,t+1} = (1-\alpha)\,A_M(c^A_{M,t})$ as well as the law of motion for q_{t+1}, we can rewrite the optimization problem as:

$$\max_{\tau^A_t(\upsilon_{M,t})} (1+\beta)(1-\alpha)\,A_M(c^A_{M,t}) - \left\{1 - \left[q_t + q_t(1-q_t)\tau^A_t(\upsilon_t)\right]\right\} \ln\left[\frac{\alpha A_M(c^A_{M,t}) s^B_t}{(1-\alpha)A_M(c^A_{M,t})}\right] - \frac{\tau^A_t(\upsilon_{M,t})^2}{2}.$$

We take the first-order condition with respect to their reaction curve $\tau^A_t(\upsilon_{M,t})$, to get:

$$\begin{aligned} \frac{\partial U^A_{M,t}(\cdot)}{\partial \tau^A_t(\upsilon_{M,t})} &= q_t(1-q_t)\tau^{A\,\prime}_t(\upsilon_{M,t}) \ln\left(\frac{\alpha}{1-\alpha} s^B_t\right) - \tau^A_t(\upsilon_{M,t})\tau^{A\,\prime}_t(\upsilon_{M,t}) = 0 \\ &\Leftrightarrow \quad \tau^A_t(\upsilon_{M,t})^* = q_t(1-q_t) \ln\left(\frac{\alpha}{1-\alpha} s^B_t\right), \end{aligned}$$

which is equation (14) in the main text.

Turning to type B agents, they move first by making a decision on the amount of their savings:

$$\begin{aligned} \max_{s^B_{M,t}} U^B_{M,t}(\cdot) &= c^B_{M,t} + \beta Y^B_{M,t+1} - E^B_t(\chi_{t+1}) - \frac{\tau^B_t(\chi_{M,t})^2}{2} \\ \text{s.t. } c^B_{M,t} + s^B_{M,t} &\leq Y^B_{M,t}. \end{aligned}$$

We first substitute out: consumption $c^B_{M,t}$ from the budget constraint; income $Y^B_{M,t+1}$ from the production function, after taking into account the marginal return to capital; the expected regime-dependent inefficiency $E^B_t(\chi_{t+1})$ from its definition after noting that $\ln \chi_{M,t+1} = 0$; under a market economy $\tau^B_t(\chi_{M,t}) = \tau^B_t(1) = 0$ (since $\chi_{M,t} = 1$, see above). Below we omit the M-subscript to savings because under communism individual savings are absent:

$$\max_{s^B_t} \left(Y^B_{M,t} - s^B_t\right) + \beta r_M s^B_t - q_{t+1} \ln(\chi_{C,t+1}).$$

Notice we take $\chi_{C,t+1}$ as given. This implies that capital owners use a notional savings rate in a market system when considering their expected loss from inefficiency. In other words, they do not increase their utility loss from a regime change by saving more by creating a higher gap between output in the two systems.

Next, we substitute: $Y^B_{M,t} = r_M s^B_{t-1}$ and $r_M = \alpha A_M(c^A_{M,t})$ (from the production function, after taking into account the marginal return to capital); q_{t+1} (from its law of motion):

$$\begin{aligned} \max_{s^B_t} \quad & r_M s^B_{t-1} - s^B_t + \beta\alpha A_M(c^A_{M,t}) s^B_t - \left[q_t + (q_t - q_t^2)\tau^A_t(\upsilon_t)\right] \ln\left(\chi_{C,t+1}\right), \\ \max_{s^B_t} \quad & \alpha A_M(c^A_{M,t}) s^B_{t-1} + \left(\beta\alpha A_M(c^A_{M,t}) - 1\right) s^B_t - \left[q_t + (q_t - q_t^2)\tau^A_t(\upsilon_t)\right] \ln\left(\chi_{C,t+1}\right). \end{aligned}$$

Replacing for $\tau^A_t(\upsilon_t)$ with the optimal reaction of type A agents from (14) we get

$$\begin{aligned} \max_{s^B_t} \quad & \alpha A_M(c^A_{M,t}) s^B_{t-1} + \left(\beta\alpha A_M(c^A_{M,t}) - 1\right) s^B_t \\ & - \left[q_t + q_t^2(1-q_t)^2 \ln\left(\frac{\alpha}{1-\alpha} s^B_t\right)\right] \ln\left(\chi_{C,t+1}\right). \end{aligned}$$

Now, we take the first-order condition to find optimal savings:

$$\begin{aligned} \frac{\partial U^B_{M,t}(\cdot)}{\partial s^B_t} &= \beta\alpha A_M(c^A_{M,t}) - 1 - \frac{q_t^2(1-q_t)^2}{s} \ln\left(\chi_{C,t+1}\right) = 0 \\ &\Leftrightarrow \quad s^{B*}_t = q_t^2(1-q_t)^2 \frac{\ln\left(\chi_{C,t+1}\right)}{\beta\alpha A_M(c^A_{M,t}) - 1}, \end{aligned}$$

where $A_M(c^A_{M,t}) > \frac{1}{\beta\alpha}$ must hold for positive savings by the capital owners. The latter expression for optimal savings is exactly equation (15) in the main text.

Substituting (15) back into (14) to derive the socialization effort in its final form, we get

$$\tau^A_t(\upsilon_{M,t})^* = q_t(1-q_t)\ln\left(\frac{\alpha q_t^2(1-q_t)^2\ln\left(\chi_{C,t+1}\right)}{(1-\alpha)[\beta\alpha A_M(c^A_{M,t})-1]}\right),$$

which is equation (16) in the main text.

Substituting $\tau^A_t(\upsilon_{M,t})^*$ from (16) into (11), next-period intensity of mobilization becomes

$$\begin{aligned} q_{t+1} &= q_t + (q_t - q_t^2)\tau^A_t(\upsilon_{M,t})^* \\ &= q_t + q_t^2(1-q_t)^2\ln\left(\frac{\alpha q_t^2(1-q_t)^2\ln\left(\chi_{C,t+1}\right)}{(1-\alpha)[\beta\alpha A_M(c^A_{M,t})-1]}\right), \end{aligned}$$

which is equation (17) in the main text.

B.2 The von Stackelberg Game under the Communist Plan

The central planner here is the first mover and decides on savings. In this aggregate optimization problem, he takes into consideration the socialization reaction of type B agents to the inefficiency caused by the centrally chosen savings. Starting with type B agents, they take savings as fixed and maximize:

$$\max_{\tau^B_t(\chi_{C,t})} U^B_{C,t}(c^B_{C,t}, Y^B_{C,t+1}, \chi_{t+1}) = c^B_{C,t} + \beta Y^B_{C,t+1} - E^B_t(\chi_{t+1}) - \frac{\tau^B_t(\chi_{C,t})^2}{2}.$$

We substitute for the expected regime-dependent inequality $E^B_t(\chi_{t+1})$ from its definition after noting that $\ln\chi_{M,t+1} = 0$; and $\chi_{C,t+1}$ (from its definition):

$$\max_{\tau^B_t(\chi_{C,t})} c^B_{C,t} + \beta Y^B_{C,t+1} - q_{t+1}\ln\chi_{C,t+1} - \frac{\tau^B_t(\chi_{C,t})^2}{2},$$

$$\max_{\tau^B_t(\chi_{C,t})} c^B_{C,t} + \beta Y^B_{C,t+1} - \left\{q_t + q_t(1-q_t)[-\tau^B_t(\chi_t)]\right\}\ln\left(\upsilon_{M,t+1}\right) - \frac{\tau^B_t(\chi_{C,t})^2}{2},$$

$$\max_{\tau^B_t(\chi_{C,t})} c^B_{C,t} + \beta Y^B_{C,t+1} - \left\{q_t + q_t(1-q_t)[-\tau^B_t(\chi_t)]\right\}\ln\left(\frac{\alpha A_M(c^A_{M,t})s^B_t}{H_{C,t}}\right) - \frac{\tau^B_t(\chi_{C,t})^2}{2},$$

$$\max_{\tau^B_t(\chi_{C,t})} c^B_{C,t} + \beta Y^B_{C,t+1} - \left\{q_t + q_t(1-q_t)[-\tau^B_t(\chi_t)]\right\}\ln\left(\frac{\alpha A_M(c^A_{M,t})s^B_t}{\alpha A_C(c_{C,t})S_t}\right) - \frac{\tau^B_t(\chi_{C,t})^2}{2}.$$

Taking the first-order condition:

$$\begin{aligned}\frac{\partial U^B_{C,t}(\cdot)}{\partial \tau^B_t(\chi_{C,t})} &= q_t(1-q_t)\tau^{B\,\prime}_t(\chi_{C,t})\ln\left(\frac{A_M(c^A_{M,t})s^B_t}{A_C(c_{C,t})S_t}\right) - \tau^B_t(\chi_{C,t})\tau^{B\,\prime}_t(\chi_{C,t}) = 0 \\ &\Leftrightarrow \quad \tau^B_t(\chi_{C,t})^* = q_t(1-q_t)\ln\left(\frac{A_M(c^A_{M,t})s^B_t}{A_C(c_{C,t})S_t}\right),\end{aligned}$$

which is equation (19) in the main text.

The central planner as a first mover maximizes utility in the name of the type A agents taking into account aggregate values.

$$\begin{aligned}\max_{S_t} \quad U^A_{C,t}(C_t, K_{C,t+1}, \upsilon_{t+1}) &= C_t + \beta K_{C,t+1} - E^A_t(\upsilon_{t+1}) - \frac{\tau^A_t(\upsilon_{C,t})^2}{2} \\ \text{s.t. } C_t + S_t &\leq H_{C,t}.\end{aligned}$$

Note that, parallel to the market economy, where $Y_{M,t+1}$ was the basis for private investment decisions, the central planner uses $K_{C,t+1} = \alpha A_C(c_{C,t})S_t$ in his optimization problem.[12] Substituting out consumption and savings from (7); the expected regime-dependent inequality $E^A_t(\upsilon_{t+1})$ from its definition after noting that $\ln \upsilon_{C,t+1} = 0$; q_{t+1} (from its law of motion); and $\tau^A_t(\upsilon_{C,t}) = \tau^A_t\left(\frac{Y^B_{C,t}}{Y^A_{C,t}}\right) = \tau^A_t(1) = 0$ under a centralized economy (since $\upsilon_{C,t} = \frac{Y^B_{C,t}}{Y^A_{C,t}} = 1$, see above).

$$\max_{S_t} \quad H_{C,t} - S_t + \beta K_{C,t+1} - (1-q_{t+1})\ln(\upsilon_{M,t+1}).$$

$$\begin{aligned}&\max_{S_t} \quad A_C(c_{C,t})[\alpha K_t + (1-\alpha)L] - (1-\beta\alpha A_C(c_{C,t}))S_t - (1-q_{t+1})\ln(\upsilon_{M,t+1}), \\ &\max_{S_t} \quad A_C(c_{C,t})[\alpha S_{t-1} + (1-\alpha)L] - (1-\beta\alpha A_C(c_{C,t}))S_t \\ &-\left(1-\left[q_t - (q_t - q_t^2)\tau^B_t(\chi_t)\right]\right)\ln(\upsilon_{M,t+1}).\end{aligned}$$

Replacing for $\tau^B_t(\chi_t)$ with the optimal reaction curve by type B agents derived in (19) we get

$$\begin{aligned}&\max_{S_t} \quad A_C(c_{C,t})[\alpha S_{t-1} + (1-\alpha)L] - (1-\beta\alpha A_C(c_{C,t}))S_t \\ &-\left(1-\left[q_t - q_t^2(1-q_t)^2\ln\left(\frac{A_M(c^A_{M,t})s^B_t}{A_C(c_{C,t})S_t}\right)\right]\right)\ln(\upsilon_{M,t+1}).\end{aligned}$$

Now, we take the first-order condition:

[12] Recall that due to the assumption of full depreciation, standard in overlapping-generations models, $K_{C,t+1} = (1-\delta)K_{C,t} + I_{C,t}$ with $\delta = 1$ and $I_{C,t} = S_t$ becomes $K_{C,t+1} = I_{C,t} = S_t$.

$$\begin{aligned}\frac{\partial U^A_{C,t}(\cdot)}{\partial S_t} &= (\beta\alpha A_C(c_{C,t}) - 1) + \frac{q_t^2(1-q_t)^2\ln(\upsilon_{M,t+1})}{S_t} = 0 \\ &\Leftrightarrow \quad S_t^* = \frac{q_t^2(1-q_t)^2\ln(\upsilon_{M,t+1})}{1-\beta\alpha A_C(c_{C,t})},\end{aligned}$$

where $A_C(c^A_{M,t}) < \frac{1}{\beta\alpha}$ must hold for positive savings by the the central planner. The latter expression for optimal savings is exactly equation (20) in the main text.

Substituting (20) back into (19) to derive the socialization effort of type B in its final form, we get

$$\tau_t^B(\chi_{C,t})^* = q_t(1-q_t)\ln\left(\frac{A_M(c^A_{M,t})s_t^B[1-\beta\alpha A_C(c_{C,t})]}{A_C(c_{C,t})q_t^2(1-q_t)^2\ln(\upsilon_{M,t+1})}\right),$$

which is equation (21) in the main text.

Substituting $\tau_t^B(\chi_{C,t})^*$ from (21) into (11), next-period proportion of the population in favor of a market-based system becomes

$$\begin{aligned}q_{t+1} &= q_t - (q_t - q_t^2)\tau_t^B(\chi_{C,t})^* \\ &= q_t - q_t^2(1-q_t)^2\ln\left(\frac{A_M(c^A_{M,t})s_t^B[1-\beta\alpha A_C(c_{C,t})]}{A_C(c_{C,t})q_t^2(1-q_t)^2\ln(\upsilon_{M,t+1})}\right),\end{aligned}$$

which is equation (22) in the main text.

C Figures

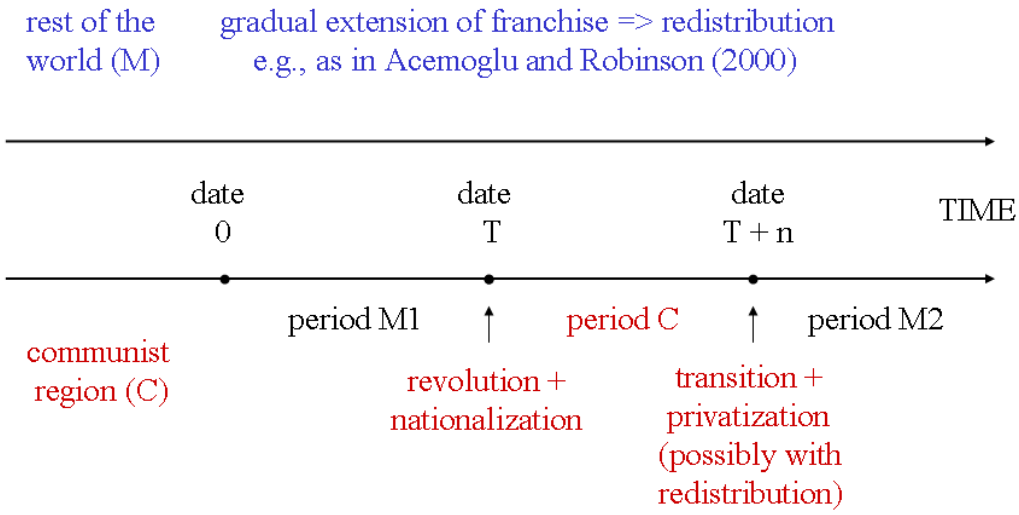

Figure 1: Model Timing and Structure

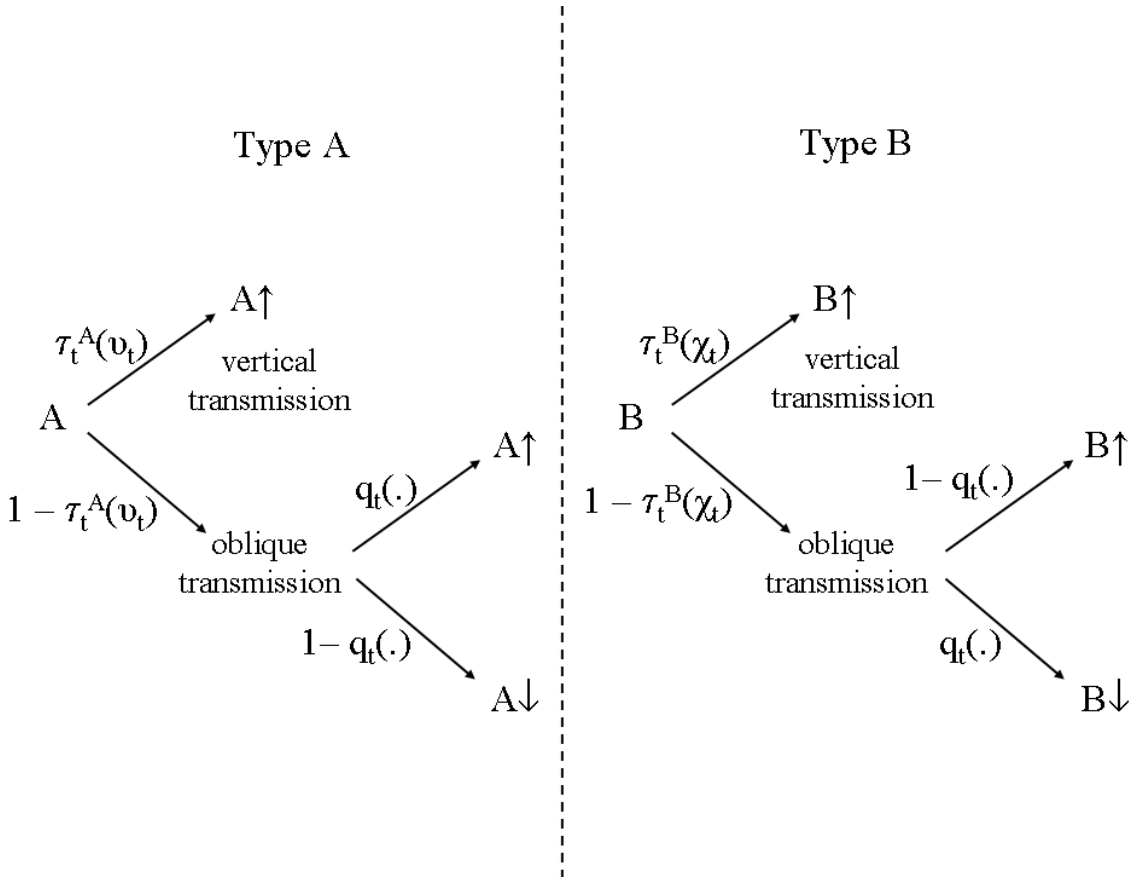

Figure 2: Types of Agents and Socialization