A Model of Sequential Reforms and Economic Convergence: Case of China*

Yong Wang†

This Version: February 2013

Abstract

Motivated by China's experience and cross-country evidence, a growth model is developed to explain the repeated interaction between reforms and convergence. International learning externality fosters convergence until a growth bottleneck is reached, at which point convergence stops unless the institution is improved. After the reform, convergence resumes until a new bottleneck is encountered, which triggers another reform, so on and so forth. Recursive method is used to characterize this dynamic interactive process. Three analytical results obtain under perfect international credit market. First, each reform occurs precisely when the new growth bottleneck just becomes binding. Second, the reform size changes monotonically over time. Third, reforms occur for a finite number of times and convergence is incessant until the last constraint binds, so a permanent GDP gap may still exist in the long run. Under imperfect credit market, advantage of backwardness may exist in economic reforms: An initially richer economy is more likely to adopt insufficient reforms because the growth bottleneck is reached too soon before enough saving has been accumulated. Moreover, reforms may be delayed resulting in intermittent convergence. The model also implies that a politically more powerful government should adopt more gradual reforms, *ceteris paribus.*

Key Words: Convergence, Growth Models, Optimal Reform, Chinese Economy, Transitions

JEL Classification Codes: O11,O43, O53, P21

*This paper is a substantial revision of the fourth chapter of my Ph.D. Dissertation at the University of Chicago. I am deeply indebted to Bob Lucas for his patient and valuable guidance. For valuable discussions, I also thank Fernando Alvarez, Francesco Caselli, Qiang Gong, Lars Peter Hansen, Justin Yifu Lin, Rachel Ngai, Steve Parente, Nancy Stokey, Francesco Trebbi, and Xiaobo Zhang. This paper was presented at University of Chicago, Peking University, Fudan University, Far East Econometric Society Meeting (2010), and Royal Economic Society Conference (2011). The usual disclaimer applies.

†Mailing Address: Department of Economics, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong. E-mail address: yongwang@ust.hk. Phone: (852)23587625. Fax: (852)23582084.

1 Introduction

The last several decades witnessed institutional transitions in China, India, Russia, Vietnam, and many other developing countries. Some were successful and managed to converge to the richest economies, but others failed. To understand why, Rodrik (2005) reviews a vast pertinent literature of reforms and economic accelerations. He finds that economic accelerations, which are strictly defined and distinct from mere recoveries from recessions, typically occurred when some institutional bottleneck was relaxed. Based on a huge amount of compelling evidence, he concludes that to ignite economic convergence in a developing economy may only need a small institutional or policy change, but to sustain the convergence it would require a process of *cumulative institutional building* along the way:

"*In the long run, the main thing that ensures convergence with the living standards of advanced countries is the acquisition of high-quality institutions. The growth-spurring strategies have to be complemented over time with a cumulative process of institution building to ensure that growth does not run out of stream...*"

This view, and the closely related view that institutions are fundamental causes of long run growth, is widely accepted by economists (see North (1990), Hall and Jones(1999), Acemoglu et al (2005)). Surprisingly, however, there exist few, if any, theoretical models that explicitly characterize how economic convergence occurs with an *endogenously cumulative process of institutional building* in the standard growth framework including the convergence literature.[1] In this paper I aim to help fill this gap.

More specifically, a multitude of case studies and anecdotal stories in developing countries suggest that there exists a regular pattern for convergence and reforms, that is, a policy or institutional reform ignites economic convergence, which continues until the economy meets a new binding constraint. Convergence stops unless the new obstacle is eliminated via a new round of reform. In other words, the catch up process triggers, and is also sustained by, endogenous and successive institutional reforms, especially in those economies that have succeeded in catching up after World War II (Wade (1990), Rodrik (2005), World Bank (2005)). In fact, this repeated interaction between convergence and reform is a hallmark feature of China's gradualist reform and growth in the past three decades (Roland (2000), Naughton (2007), Rodrik (2010), Xu(2011), Lin (2012)), which is reviewed in detail in Section 2. In this paper, I will develop a theoretical model to precisely capture this interactive process between economic growth (convergence) and sequential relaxations of newly-arising constraints, which is so far only informally described

[1] Acemoglu et al (2006) show that, in order to achieve sustained development, the growth mode should switch from the investment-based strategy to innovation-based strategy when a country gets closer to the world technology frontier. However, they do not emphasize the "cumulative process" of sequentially relaxing the new institutional constraints.

in the literature. My focus is on the analytical characterization of their mechanical interactions. I discipline the modeling by making minimum deviations from standard growth models.

Institutional bottlenecks are different in nature and change over time, but, for our purpose, all these binding constraints and the related reforms must be addressed simultaneously in a unified framework. To maintain tractability, I put forth a highly stylized model and address the following two normative and simple questions: what is the first-best reform scheme and how does it interact with economic convergence? We must have confident answers to these benchmark questions before any further sophisticated analysis can be made. Following Lucas (2009), I study these two questions in a framework of human capital externality (idea flows) mainly for its simplicity. It is intrinsically related to the question how to eliminate barriers to adopting foreign better technology, which is explored by Parente and Prescott (1994, 2000).[2]

In the existing growth literature, economic convergence is usually studied with institutions and policy barriers taken as exogenous and time-invariant (Barro and Sala-I-Martin (1992), Ngai (2004), Stokey (2012)). My approach is different in that the barrier will be an endogenous policy variable instead of an exogenous parameter. Economic reforms are represented by endogenous reduction of this barrier variable in my model. Same as Lucas (2009), a developing economy grows as its human capital accumulates thanks to the learning externality (spillover) from a representative and exogenously growing developed economy. Their initial large gap in human capital (and GDP) allows the poor country to catch up. In my model, the new element is that now convergence stops when their gap shrinks to a threshold value, which depends on the barrier variable. This convergence bottleneck is referred to as the learning constraint being binding. The benevolent social planner, or Ramsey government, formulates a barrier adjustment scheme, which specifies when and how much to change this barrier variable to maximize the social welfare. Using the recursive method, I analytically characterize a non-stationary dynamic optimization problem to show how economic growth triggers barrier reduction (institutional reforms) and how the latter feeds back on factor accumulation and growth dynamics.

Four main results are found. First, each reform occurs precisely when the learning constraint just becomes binding. Second, the magnitude and frequency of reforms are both monotonic over time. Moreover, if the size of barrier reduction becomes smaller (larger) over time, the time interval between two adjacent reforms becomes shorter (longer) over time. Third, the number of reforms is finite and the successive reforms support an incessant convergence until the learning constraint binds again after the last reform. In the long run, a GDP gap may still exist but the two countries have the same growth rate. Fourth, when the international credit market is imperfect, reforms may be delayed and the resulting convergence process can be punctuated and intermittent. In particular,

[2] Klenow and Rodriguez-Clare (2005) provides an excellent survey for human capital learning externality and our paper is also closely related to discussions on human capital and technology diffusion by Benhabib and Spiegel (2005). Stokey (2012) examines explicitly the different roles played by human capital and technology diffusion in the catching up process.

there may exist an advantage of backwardness in reforms. Under certain conditions, an initially richer economy is more likely to have an insufficient reform because growth bottleneck is reached too soon before enough saving is accumulated. Following Parente and Prescott (2000), I place more emphases on level effect than growth effect in the analysis.[3]

My model highlights the importance of cumulative and sequential reforms that underlie the entire process of economic convergence. At a superficial level, the course of convergence still appears to be fully dictated by the human capital accumulation without any explicit role for the barrier variable, seemingly identical to the existing convergence literature with exogenous barriers. However, at a deeper level, my model mechanism strongly echoes the view of Acemoglu et al (2005) and Rodrik (2005) by demonstrating explicitly that, beneath the GDP dynamics, the fundamental force that ignites and sustains economic convergence is actually the successive reforms in institutions or policies, in a form of sequentially eliminating newly-arising binding constraints *along* the process of convergence, the empirical relevance of which is repeatedly established and stressed in the growth and transition literature (Qian (2003), Lin (2012), World Bank (2005), Hausmann et al (2008) and Rodrik (2005, 2010)). Unfortunately, so far few attempts, if any, have been made to formalize this important idea in the pertinent growth and reform literature.

My model also sheds light on optimal reform strategies. One implication is about the optimal timing of reforms: reforrms should be made at the point of crises, which in my model refers to learning constraint being binding because further convergence is jeopardized at that moment. This prediction is consistent with the empirically-verified "crisis hypothesis" in the reform literature, but the mechanism is different from the political economy argument (Drazen and Grilli (1993), Alesina et al (2006)). My model also implies that the optimal reform scheme depends on whether the administrative power of the government, as it determines the relative importance of the fixed and variable reform costs. A strong government, which may be authoritarian, tends to formulate and implement a reform plan more quickly than a weak government, so the fixed cost for each reform is smaller, *ceteris paribus*. The opposite can be true for the variable cost of reform, because a weak government is more likely to have already negotiated with different parties to minimize the negative impact on different groups for any given reform size.[4]This intrinsic asymmetry, according to the model, implies that the economy with a stronger government *should* have a more gradual reform while one with a weak government should have a more radical reform. These results are diametrically opposite to the standard "Washington Consensus", which recommends that all imperfect institutions must be reformed thoroughly and simultaneously at the fastest speed regardless of the

[3] In growth models with human capital learning externality (without institutional change), a standard result is that the developing and the developed countries grow at the same speed on the balanced growth path, as implied by the law of motion in the human capital diffusion (Klenow and Rodriguez-Clare, 2005). But the level difference can be still enormous. Parente and Prescott (2000) argue that it is important to look at the level effect too.

[4] A recent growing literature emphasizes the importance of state capacity and the shortcomings of a weak government in economic development, see Blanchard and Shleifer (2001), Acemoglu (2005), Besley and Persson (2009).

cost structures of reforms or the strength of government. This therapy predicts that economic convergence does not take place until after the thorough reform and then no binding constraint will ever occur as an economy grows.[5]

How capital investment is affected by fixed and variable adjustment costs has been intensively discussed in the textbook investment model with convex adjustment cost functions and in the more sophisticated (S, s) inventory models. However, methodologically, a distinctive feature of my model is that it studies the "capacity adjustment" instead of "flow adjustment". Recall that (S, s) inventory models examine how to adjust the inventory "flow" under the constraint that the "container" has an exogenous capacity, but what is optimally adjusted in this paper is the capacity of the "container" rather than the "flow" itself. Notice that my model does not directly adjust human capital as it cannot jump discontinuously, which is different from physical capital investment in the pertinent investment literature.

In this regard, Parente (1994) is similar to the (S, s) model because the choice variable is the technology level per se instead of the capacity or ability of technology adoption. A more important difference between Parente's model and mine is the budget constraint. In his model the adjustment cost is the loss of expertise thus the cost is not paid explicitly with physical resources, hence the budget constraint problem is circumvented. In contrast, reforms require explicit financial resources in my model and I show that different degrees of international credit market perfection, which affects the budget constraint of the reformer, have significant impact on reform strategies and growth dynamics.

The existing academic literature of reforms and transitions, except for literature reviews, typically focuses on one, or at most two, specific institution or policy change(s) at each time. The merit of this approach is that it allows for explicitly incorporating more institutional details into the analysis,[6] but this separate treatment on different reforms does not help organize our thinking about the interactive process of different reforms and convergence at different development stages, as we observe in China and many other economies. It is this interactive process that I try to formalize and shed light on. Dewatripont and Roland (1992), Wei (1997), Lau et al (2000), Caselli and Gennaioli (2008), among others, study the optimal sequencing of reforms by analyzing how to make it politically feasible to gain enough support to push forward different reforms in the presence of multiple groups with conflicting interests, but these models are not formulated in an explicit growth framework, so it is not clear how reforms and convergence interact with each other. Moreover, they focus on the role of political constraints rather than

[5] The "Washington Consensus" has received wide criticisms as it is incapable of explaining many empirical facts including, for example, the stagnation of the average developing countries in 1980-1998, although they all followed the suggestions prescriptions by the "Washington Consensus" (Easterly (2001)). The skepticism reached its peak after the former Soviet Union and some other East European countries experienced unexpectedly huge economic difficulties after adopting the shock therapy. In sharp contrast, China obtained unexpected success by adopting a more pragmatic and piece-meal reform strategy (see Stiglitz (1998), Rodrik (2005, 2010), World Bank(2005), Xu (2011), Lin (2012).

[6] See, for example, Murphy et al (1992), Lin (1992), Barberis et al (1996), Banerjee et al (2002), Easterly (2005), and Wang (2013). Excellent surveys include Xu (2011) and Roland (2000).

the sequentially binding growth bottlenecks. My paper complements this literature as the contribution of this paper is not to provide new insights on some specific reform, but rather to help bridge the gap between the two different, often orthogonal, strands of literature: economic convergence and sequential reforms.

The rest of the paper is organized as follows: Section 2 provides two specific motivating stories from China. Section 3 describes the general setting of the model economy. Section 4 and Section 5 examine the optimal reform and the associated growth dynamics when the international credit market is perfect and imperfect, respectively. Section 6 concludes by discussing some possible avenues for future research.

2 Motivating Stories from China

To help motivate my investigation and to help theorize the general interactive process of barrier reduction (policy change or institutional building) and economic convergence (economic growth), it is useful to start by looking at two concrete stories from China.

One story is about the development process of a potato industry in Anding county in Northwest China. Before the Household Responsibility System (HRS) was adopted in the late 1970s, all the farmers worked for the commune and were equally paid regardless of individual effort. Naturally the output was low and people could not even feed themselves. After the agricultural reform, HRS was implemented, which allowed individual households to claim their residuals after fulfilling some fixed quota. This reform solved the incentive problem, which was the main institutional bottleneck for growth at that stage, thus output grew rapidly afterwards. However, the next growth hurdle appeared soon: the land was too infertile to raise the output of traditional crops such as wheat or sorghum. To solve this problem, local farmers and the local government invited some agricultural experts from Beijing to seek solutions. It was soon discovered that some types of potatoes were magically suitable for the land conditions of that region. Those potatoes were introduced into Anding, together with the information about how to grow them efficiently. Aggregate agricultural output increased tremendously afterwards. Farmers could not only feed themselves but also have surplus potatoes to sell.

Now came the new growth bottleneck: local farmers had no price information. Consequently, farmers could only sell their surplus potatoes at low prices to a handful of local intermediaries with monopoly power. To overcome this bottleneck, some farmers jointly hired a person to physically stay in Zhengzhou, one of the major national potato markets and far away from Anding, collecting and reporting the price information, and later the county government of Anding established a new office located in Zhengzhou helping disseminate the potato price information freely to all the farmers in Anding. As information hurdle was dismantled, local farmers had good bargains for their surplus potatoes and therefore the output skyrocketed. However, the fast output growth soon made the transportation capacity become the new growth bottleneck: long-way transportation had to rely on trains, but Anding was unfortunately allocated with only two carts. After some personal contact and negotiations between the Anding government and China's Ministry of Railways, four more carts were added to the trains and the transportation constraint was effectively relaxed, after which potato output continued to climb. After witnessing the success of Anding's potato industry, the neighboring counties started to mimic the practice of Anding So even six carts soon became insufficient. Then local people cooperated with the local government and found an innovative way to hoard potatoes during the harvest season and also tried to produce and export higher value-added processed products of potatoes rather than raw potatoes, which naturally led to further growth in the local GDP of Anding (see Zhang and Hu (2011)).

In this example, we can see clearly that the industrial development encountered distinctive bottlenecks at different stages: production incentives, technology constraint, information hurdle, transportation capacity, hoarding technology, and so on. Continuous growth became admissible only when the new binding constraints were dismantled in time. Moreover, the constraints that became binding later would not have become binding if the economy had not developed after the earlier growth-binding barriers were eliminated. In other words, economic growth triggered the arrival of new binding constraints and hence called for further institutional buildings (reforms). Meanwhile, institutional buildings relaxed the growth bottlenecks and sustained economic growth. This interactive process of economic convergence and successive removal of binding obstacles is a general pattern in industrial development. Although, this is only a story about one particular industry in a particular region, it vividly demonstrates how sequential reforms and economic growth interact and stimulate each other.

The second motivating story is at the aggregate level, or virtually a coarse sketch of the whole development experience of China in the last three decades.[7] The remarkable acceleration of China's economy started by the rural reform in the late 1970s, when farmers' incentives to work were strengthened after the universal adoption of the Household Responsibility System . Rural productivity and total output increased dramatically after the incentive constraint was relaxed (Lin (1992)). However, economic growth soon led to a new bottleneck: more and more farmers were released from the agricultural sector and further growth required that more industrial jobs be created for them. However, it was unconstitutional to establish private firms at that time and it was politically infeasible to allow these abundant labor to move into cities. This binding constraint triggered the birth of the newly-innovated semi-public Township-and-Village enterprises, which were ideologically acceptable. Without fundamentally changing the constitution or law enforcement, this institutional reform facilitated rural industrialization by absorbing surplus rural labor and significantly contributed to China's further growth (Qian (2003)). After the Tiananmen event in 1989, China was sanctioned and hence cut off from the international capital market or external aid, and the economy stagnated. With limited foreign aid for its institutional adjustment, China mainly relied on its domestic savings to finance reforms.

[7] A complete narrative about China's economic development is beyond the scope of my paper and many other important features of Chinese economy are not highlighted here due to the space constraint. For more detailed treatment, readers are referred to Qian (2003), Naughton (2007), Rodrik (2010), Xu (2011), Lin (2012), and the references cited there.

The market-oriented and open-door policies resumed after Deng Xiaoping's South Tour in 1992. More and more special economic zones and joint ventures were established. Meanwhile, China managed to attract a substantial amount of FDI, which tremendously released its financial and technological constraints, and, therefore, spurred the catch-up growth even further (Wang (2013)). All the economic development paved the way for the reform of inefficient State-Owned Enterprises (SOEs), which was then the major growth obstacle for the urban areas and for the whole economy at large. Massive SOE reforms started in the late 1990s. However, instead of privatizing all the SOEs overnight like Russia, China kept the large SOEs, especially those in the upstream industries (such as energy, raw materials, telecommunications and banks), and let go small-and-medium-sized SOEs, which were mostly concentrated in the downstream industries (such as labor-intensive manufacturing and consumption-oriented service such as hotel and restaurant). Deregulation of downstream industries led to more entry of private firms and enhanced competition, resulting in efficient resource reallocation from bankrupt SOEs to more productive private firms, which enabled China to continue the high growth even though the whole financial sector was still inefficient and dominated by large State-Owned banks (Song et al (2011)).

China's entry into WTO in 2001, together with relaxations of rural-urban migration restrictions, facilitated China's industrialization and the emergence of state capitalism featured by a vertical structure: As SOEs have largely exited from the downstream industries so the downstream operates under capitalism with free entry, private ownership, and perfect competition, whereas the upstream industries are still monopolized by SOEs, which extract monopoly rents from the downstream industries by charging price markups on the intermediate goods and services needed in the downstream productions. Consequently, the profitability of SOEs exceeded that of non-SOEs in the last decade as upstream SOEs benefit disproportionately more from the expansion of non-SOEs in the downstream industries, thanks to cheap labor and downstream non-SOEs' accessibility to the world export market. However, as the economy continues to grow, labor cost starts to rise after sufficient industrialization and urbanization (China has passed the so-called Lewis turning point). It means that eventually the upstream SOE monopoly will strangle the growth of downstream private sectors because downstream firms have to pay the expensive inputs monopolized by the upstream SOEs but the advantage in cheap labor disappears, not to speak of competitions from other developing countries with cheaper labor. To sustain further growth, the government will have to reform the upstream SOEs as they are becoming the news binding constraint (Li, et al (2012)).

These two stories in China's economic development are both featured by the interactive process of successive reforms and economic convergence. These experiences are by no means unique to the Chinese economy. Wade (1990), Cander (2006), Rodrik (2005), and World Bank (2005) all provide convincing cross-country case studies and empirical evidence showing that mild policy (or institutional) changes sometimes can activate an industry or even a whole economy and the industry or economy keeps developing as sequentially-binding growth bottlenecks are eliminated one by one via policy changes or institutional reforms. Complementary to this view, Hausmann et al (2008) and Rodrik

(2010) further advocate the approach of growth diagnostics and they find that the binding institutional constraints for growth are not only different across different countries but also varying at different development stages for the same country.

3 Model Environment

Although the two detailed motivating stories are both from China, the logic of the model is meant to be more general. Consider a developing economy populated by a unit mass of identical households. A representative household maximizes the total present value of discounted utility:

$$\int_0^\infty c(t)e^{-\rho t}dt, \tag{1}$$

where ρ is the discount rate. The assumption of infinite inter-temporal elasticity of substitution helps us focus on the institutional change problem by making the consumption analysis trivial.[8] Following Lucas (2009), I assume that a representative household is endowed with one unit of labor, which is inelastically supplied to produce one homogeneous good with the following technology:

$$f(h, G) = mh^\alpha G^{1-\alpha},$$

where h is human capital and G represents the public infrastructure provided and maintained by the government. Suppose G is financed by the tax revenues on the output at rate τ, then $G = \tau f(h, G)$, which implies

$$f(h, G(h, \tau)) = \tau^{\frac{1-\alpha}{\alpha}} m^{\frac{1}{\alpha}} h,$$

Without loss of generality, we can normalize m such that the above equation is reduced to

$$\widetilde{f}(h) = h,$$

that is, one unit of human capital ultimately can produce one unit of final good, which is storable and can be either consumed or used to pay the cost of institution adjustment.[9] The initial human capital is h_0.

[8] With a non-degenerate CRRA utility function, the reform analysis becomes over complicated without generating sufficiently useful new insights because this paper wants to emphasize the level effect instead of the growth effect and the focus is on institutional reforms rather than consumtpion behaviors. An alternative way to interpret this deterministic setting is that the Ramsey government decomposes his dynamic decision into two separate steps: first, it tries to maximize the representative household's total life-time income ($\rho = r$ is assumed throughout this paper) via optimal reforms; second, dynamic consumption decisions will be made subject to the intertemporal budget constraint as identical final good can be also traded internationally at the unit price.

[9] Here human capital actually might refer to the combination of all the intangible accumulative production factors including technology. It is straightforward to introduce endogenous time allocation decision on human capital accumulation in a textbook way, but this complication does not yield any new insights for the current purpose, so I choose to abstract it away.

There is also a developed economy with the same one-to-one production technology. $H(t)$ denotes its per capita human capital stock at time t, which grows exogenously at a constant exponential speed g_H. $H(0)$ is normalized to unity. Due to the positive learning externality in human capital (or idea flows), $h(t)$ increases up to a limit determined by an institutional barrier variable, $\delta(t)$. Following Parente and Prescott (1994) and Stokey (2012), variable $\delta(t)$ captures all the institutional factors that affect the diffusion process of external human capital (or technology and ideas) at time t such as trade barriers, FDI policies, or intellectual property rights protections. A larger $\delta(t)$ means worse institutions. δ_0 denotes the initial barrier value. The law of motion for $h(t)$, and hence the GDP of the developing country, is given by

$$\dot{h}(t) = \mu h(t) + \Phi(\frac{h(t)}{H(t)}, \delta(t)) \cdot \dot{H}(t),$$

where μ is a positive parameter dictating the growth rate of domestic human capital net of depreciation in autarky; $\Phi(\frac{h(t)}{H(t)}, \delta(t))$ captures the strength of the international learning externality (knowledge spillover), which depends on the gap in human capital $\frac{h(t)}{H(t)}$ and the barrier variable $\delta(t)$. In particular, I assume

$$\Phi(\frac{h(t)}{H(t)}, \delta(t)) \equiv \begin{cases} \frac{h(t)}{H(t)}, & if \ \frac{h(t)}{H(t)} < \frac{\eta}{\delta(t)} \\ 0, & otherwise \end{cases},$$

which says that a larger gap in human capital (a smaller $\frac{h(t)}{H(t)}$) generates a stronger externality because the foreign pool of ideas to tap from is larger from the developing country's point of view. where μ and η are both positive parameters. η is a parameter useful for comparative statics analysis. A higher η implies a longer time to enjoy the convergence for any given institution barrier and gap in per capita GDP. One possible interpretation for η could be the population ratio of the developing economy relative to the developed economy, which captures the scale effect. However, the positive externality is conditional on that the learning constraint , $\frac{h(t)}{H(t)} < \frac{\eta}{\delta(t)}$, is not binding.[10] Define $x(t) \equiv \frac{h(t)}{H(t)}$, then the above two equations yield

$$\frac{dx(t)/dt}{x(t)} = \begin{cases} \mu, & if \ \ x(t) \leq \frac{\eta}{\delta(t)} \\ 0, & otherwise \end{cases} \quad \text{and } x(0) = h_0. \tag{2}$$

So the gap between the two countries shrinks at a constant exponential speed μ until the gap hits the critical value $\frac{\eta}{\delta(t)}$, at which point convergence stops unless the institutional barrier variable $\delta(t)$ is adjusted downward. This is what I mean by "institutional improvement" or "reform".

[10] This functional form is adopted for simplicity. However, it is not difficult to generalize it to a multi-value step function (or a continuous function in the limit) so that the convergence speed also depends on the gap. By construction the developing economy never grows at a speed lower than the developed economy in the model. This possibility can be incorporated by letting the relevant value on the right hand side of (2) be negative.

The reform cost has two components: a variable cost and a fixed cost. More precisely, when the barrier variable $\delta(t)$ is adjusted from δ to δ' in a single step, the cost is given by

$$C(\delta,\delta') = \begin{cases} A\left(\frac{\delta}{\delta'}\right)^{\phi} + B, & if\ \delta \neq \delta' \text{ and } \delta' \geq \eta \\ \infty, & if\ \delta' < \eta \\ 0, & if\ \delta = \delta' \end{cases}, \tag{3}$$

where parameters A and B are both positive. $\phi > 1$, so the adjustment cost function is convex in the adjustment size, $\frac{\delta}{\delta'}$. No adjustment ($\delta = \delta'$) naturally incurs no cost. (3) also imposes a lower bound for η, which is to rule out the leapfrogging of the developing economy by only exploiting international human capital externality or automatic technology diffusion. So $x(t) \leq 1, \forall t$.

The reform cost depends on the structural details of the political institutions, so (3) should be interpreted as a reduced form for the overall cost associated with reforms. A bigger reform is more costly, captured by the variable cost, $A\left(\frac{\delta}{\delta'}\right)^{\phi}$. Given the size of institution adjustment, countries that implement economic reforms mainly through administrative orders and centralized planning are more likely to create larger distortions and hence incur a higher social cost, so A would be larger, as compared with the pro-market reform strategies in a more deregulated economy. On the other hand, the fixed cost B may include all the opportunity costs of proposing a reform plan and getting it passed in the legislature. B is large if intensive multilateral bargaining and negotiations are always involved in each reform process. The more powerful and politically consolidated the central government, the smaller the fixed cost B.[11] Dixit (2004) explicitly discusses costs of institutional building, and he argues that setting up formal institutions (such as legal rules and democratic political systems) requires high fixed costs B but low marginal (variable) costs A, whereas informal institutions (such as moral codes and common practice) are the opposite.

Reform reversals (upward adjustment of δ) are allowed, but a benevolent government has no incentive to do so, and hence the relevant adjustment must be downward. The functional space for the reform policy function is

$$\Delta \equiv \{\text{real function } \delta(t) : \mathbb{R}_+ \to [\eta, \delta_0] \text{ such that } \delta(0) = \delta_0\}.$$

Since B is positive, conventional reasoning implies that $\delta(t)$ must be a step function due to the discontinuity of adjustments. More precisely, the Ramsey government needs to find a bounded and weakly decreasing sequence $\{\delta_i\}_{i=0}^{\infty}$ and the corresponding adjustment time sequence $\{t_i\}_{i=0}^{\infty}$ with given δ_0 and $t_0 = 0$, where δ_i and t_i stand for, respectively, the value of the barrier variable right after the ith adjustment and the time of that adjustment.

[11] Blanchard and Shliefer (2001) argue that one important reason why the decentralization economic reform was successful in China but failed in Russia in the 1990s is because China was more politically centralized and hence every step of the reform was under control by the strong central government, whereas Russian central government at that time was too weak to maintain orders or implement effective reforms, thus the reforms turned chaotic.

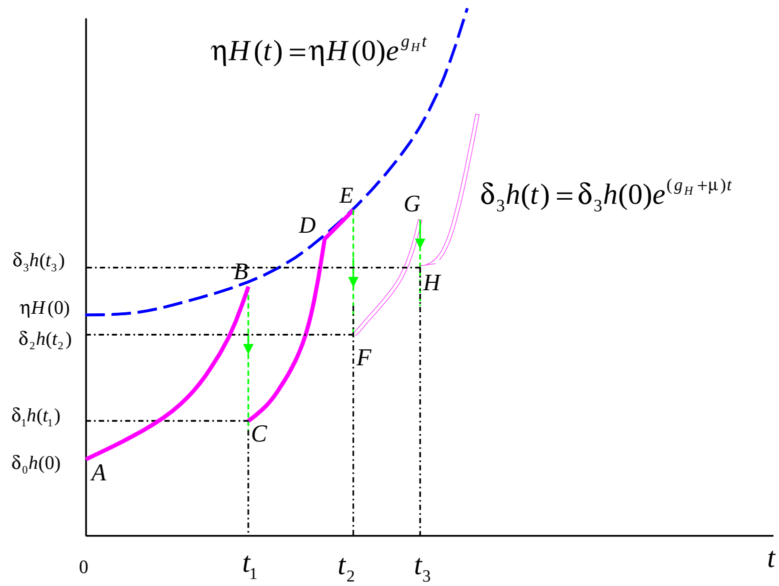

Figure 1. A Possible Adjustment Scheme

Figure 1 depicts what a possible (not necessarily optimal) adjustment path may look like. The dashed curve plots $\eta H(t)$, which grows at the exponential speed g_H. The solid curve is $h(t)\delta(t)$, which is the developing economy's human capital stock multiplied by the barrier variable $\delta(t)$. At time 0, the developing economy is at point A.No institutional adjustment is made so (2) implies that $h(t)$ grows at the exponential speed $g_H + \mu$ until time t_1, when the solid line hits the dashed line at point B. That is, the learning constraint becomes binding. The barrier variable is adjusted downward from δ_0 to δ_1 at time t_1, so $\delta(t)h(t)$ jumps down to point C. Note that human capital cannot jump. The adjustment cost $C(\delta_0, \delta_1)$ is paid at t_1. After this reform, the learning constraint is relaxed so $h(t)$ continues to grow at speed $g_H + \mu$ until the learning constraint becomes binding again at point D. Since no reform is made, convergence stops and $h(t)$ can only grow at speed g_H afterwards.[12] So the solid curve overlaps with the dashed curve. The second reform is made at time t_2, at which point the developing economy jumps from point E to point F due to the downward change of the barrier variable to δ_2. The convergence resumes. At time t_3, the learning constraint is not binding yet, but the third reform may be chosen to implement at this time point, so the economy jumps from point G to point H, so on and so forth. My task is to find the optimal adjustment scheme, namely, the optimal solid curve such that the representative household's goal function (1) is maximized.

Before I mathematically characterize this dynamic reform problem, which appears to be mechanical, it may be important to highlight its economic relevance. As can be seen from the motivating example in Section 2, reform and growth (convergence) interact each other repeatedly. Reform is needed to sustain the catch-up process, which in turn leads to sequentially binding growth bottlenecks as the economy grows. Binding constraints are different at different development stages. These newly-arising bottlenecks then trigger

[12] It is consistent with the standard result that, in the long run equilibrium, the developed and developing countries have the same grow rate on the balance growth paths, as in Klenow and Rodriguez-Clare (2005) and Benhabib and Spiegel (2005).

further rounds of reforms. Without catch-up growth, new institutional bottlenecks would not present themselves in the first place. This is precisely the logic behind China's pragmatic approach of the gradual reforms. Figure 1 shows that institutions can be improved successively along with economic growth. Alternatively, δ_0 can be thoroughly improved to the perfect level η in one step (if financially feasible), and then the economy enjoys convergence without the necessity to conduct reforms in the future. Which one is better and what else do we know? These issues will be exactly addressed by this model.

(2) implies that it can be assumed without loss of generality that $g_H = 0$ as I am most interested in convergence (relative performance) of the developing country.[13] Thus, $x(t) \equiv h(t), \forall t$. The interest rate r in the developing country is exogenously determined by the international credit market. I set r equal to ρ. To make the analysis empirically relevant and theoretically concise, I focus on the case when $\mu > r$.[14] Section 4 studies the problem when the international credit market is perfect so that the developing economy can borrow internationally. Optimal Reform under imperfect credit market is studied in Section 5.

4 Reform Under Perfect Credit Market

When international borrowing is allowed, any optimal (hence beneficial) institutional adjustment by definition must satisfy the budget constraint. The Ramsey government needs to find an optimal adjustment scheme, $\{\delta_i, t_i\}_{i=1}^{\infty}$, and an optimal time path of consumption, $c(t) \geq 0, \forall t$, to maximize (1) subject to (2), (3), with δ_0 and h_0 given, and subject to the following budget constraint:

$$\int_0^{\infty} c(t)e^{-rt}dt \leq \sum_{i=0}^{\infty}\left[\int_{t_i}^{t_{i+1}} h(t)e^{-rt}dt - C(\delta_i, \delta_{i+1})e^{-rt_{i+1}}\right], \tag{4}$$

that is, the total present value of consumption must not exceed the total present value of output (income) net of all the reform costs.[15] t_0 is set equal to 0. The status quo is maintained if the net benefit of the reform is zero.

[13] To see this, we can simply define $\widehat{\mu} \equiv \mu + g_H$ when $g_H > 0$ with $\widehat{\mu}$ reinterpreted as the catching-up speed or relative speed between the two economies. So learning externality still exists when the learning constraint binds, even though convergence stops.

[14] μ is the difference in the growth rates between the two countries during the convergence process. When the two countries are China and US, μ is clearly larger than the annual interest rate of the risk-free treasury bills in the last thirty years. Theoretically, it is straightforward to analyze the case when $\mu \leq r$ by following exactly the same method, but no additional insights can be obtained.

[15] The model is cast as a central planner problem rather than a competitive equilibrium problem for reasons beyond the second welfare theorem and modeling convenience: (1) some important markets may be missing in the less developed economy, hence resource allocation may not fully operate through the market mechanism, and (2) in reality the central governments in many transitional economies have a far greater administrative power than their counterparts in the developed economies, both in terms of shaping and changing the institutions. In reality, in many developing countries such as China or India, the central governments do have and implement very formal and extensive five-year, ten-year or twenty-year plans to reform economic institutions.

Given δ_0 and h_0, the social planner's problem can be rewritten as follows:

$$V(\delta_0, h_0) \equiv \max_{\{\delta_i, t_i\}_{i=1}^{\infty}} \sum_{i=0}^{\infty} [\int_{t_i}^{t_{i+1}} h(t) e^{-rt} dt - C(\delta_i, \delta_{i+1}) e^{-rt_{i+1}}], \tag{5}$$

subject to (2), (3), and that the associated adjustments must be always affordable:

$$\sum_{i=0}^{\infty} [\int_{t_i}^{t_{i+1}} h(t) e^{-rt} dt - C(\delta_i, \delta_{i+1}) e^{-rt_{i+1}}] \geq 0. \tag{6}$$

The key analytical challenge lies in the fact that the optimization problem is non-stationary in the sense that there exists no fixed point for the value function or the implicit policy function for the associated Bellman equation. This is mainly because of the discontinuity of the catch-up speed before and after learning constraints become binding (see (2)) and that $\delta(t)$ may change discontinuously for only a finite number of times (to be clear soon). However, this problem can still be analyzed recursively. Let N denote the total number of adjustment opportunities that are available to the planner. I first set N to be a given finite number and examine the corresponding mechanics of this dynamic system. Let V_N denote the value function with a total of N adjustment opportunities. Later, I will set $N = \infty$ and explore the optimal number of adjustment options that are actually needed. Observe that $V(\delta_0, h_0)$ in (5) must be bounded both from above and from below because $h(t) \leq 1$, $\forall t$.

4.1 No Adjustment Opportunity ($N = 0$)

When $N = 0$, convergence occurs until the learning constraint becomes binding (at point B shown in Figure 1), so GDP evolves as follows

$$h(t) = \begin{cases} h_0 e^{\mu t}, & \text{if } t < \widehat{t} \\ h_0 e^{\mu \widehat{t}}, & \text{if } t \geq \widehat{t} \end{cases},$$

where $\widehat{t}$ is the time point when the learning constraint just binds:

$$\widehat{t} = \max\{0, \frac{1}{\mu} \ln \frac{\eta}{\delta_0 h_0}\}. \tag{7}$$

The corresponding value function with zero adjustment is given by

$$V_0(\delta_0, h_0) = h_0 \int_0^{\widehat{t}} e^{\mu t} e^{-rt} dt + e^{-r\widehat{t}} \int_0^{\infty} \frac{\eta}{\delta_0} e^{-rt} dt,$$

which, by revoking (7), yields

$$V_0(\delta_0, h_0) = \begin{cases} \frac{\mu h_0}{r(\mu - r)} \left(\frac{\eta}{\delta_0 h_0} \right)^{\frac{\mu - r}{\mu}} - \frac{h_0}{\mu - r}, & \text{if } \delta_0 h_0 < \eta \\ \frac{h_0}{r}, & \text{if } \delta_0 h_0 \geq \eta \end{cases}. \tag{8}$$

To avoid analytical triviality, the initial income gap is assumed to be sufficiently large so that reforms are desirable:

Assumption A0:

$$h_0 < \eta/\delta_0. \tag{A0}$$

4.2 One Adjustment Opportunity ($N = 1$)

Let t_1 denote the time when the barrier variable is adjusted. The control can be exercised either weakly before or weakly after the learning constraint binds, so the value function is given by

$$V_1(\delta_0, h_0) = \max\{G_1(\delta_0, h_0), F_1(\delta_0, h_0)\},$$

where

$$G_1(\delta_0, h_0) \equiv \max_{t_1 \le \widehat{t}, \delta_1 \ge \eta} \int_0^{t_1} h_0 e^{\mu t} e^{-rt} dt + e^{-rt_1} \left[V_0(\delta_1, h_0 e^{\mu t_1}) - C(\delta_0, \delta_1) \right], \tag{9}$$

and

$$F_1(\delta_0, h_0) \equiv \max_{\widehat{t} \le t_1, \delta_1 \ge \eta} \left[\begin{array}{c} \int_0^{\widehat{t}} h_0 e^{\mu t} e^{-rt} dt + \int_{\widehat{t}}^{t_1} \frac{\eta}{\delta_0} e^{-rt} dt \\ +e^{-rt_1} \left[V_0(\delta_1, \frac{\eta}{\delta_0}) - C(\delta_0, \delta_1) \right] \end{array} \right] \tag{10}$$

$$= \max_{\widehat{t} \le t_1, \delta_1 \ge \eta} \left[\begin{array}{c} V_0(\delta_0, h_0) + \\ +e^{-rt_1} \left[V_0(\delta_1, \frac{\eta}{\delta_0}) - C(\delta_0, \delta_1) - V_0(\delta_0, \frac{\eta}{\delta_0}) \right] \end{array} \right]. \tag{11}$$

The following lemma says that the optimal reform time is weakly after the learning constraint first gets binding.

Lemma 1. $V_1(\delta_0, h_0) = F_1(\delta_0, h_0)$ *for any* (δ_0, h_0) *that satisfies Assumption A0.*

Proof. Refer to Appendix 1. ■

The intuition is straightforward. For any adjustment made strictly before the learning constraint becomes binding, the net value can be strictly increased if the same size of adjustment is made at $\widehat{t}$. This is because the gross benefit of any such adjustment is independent of the adjustment time before the learning constraint binds and the same adjustment cost is now paid later. This lemma allows us to focus on the adjustment made only weakly after the learning barrier becomes binding.

Lemma 2. $t_1^* = \widehat{t}$ *if* $\delta_1^* < \delta_0$ *and* $t_1^* < \infty$.

Proof. Refer to Appendix 2. ■

Lemma 2 states that the barrier adjustment, if made, must occur when the learning constraint just binds. The intuition is that any further delay is costly without adding any benefit because the optimal adjustment size remains the same, as indicated by the second term in equation (11). The optimal adjustment size is obtained from the first order condition. It is an interior solution if and only if the following is true:

$$A\phi r \leq \frac{\eta}{\delta_0} \leq (A\phi r)^{\frac{1}{\phi+\frac{r}{\mu}}}, \tag{12}$$

where the first inequality ensures that the adjustment is downward ($\delta_1 \leq \delta_0$) and the second inequality ensures that the new barrier is no smaller than η (that is, $\delta_1 \geq \eta$). For the convenience of exposition, define

$$\widetilde{B}(z) \equiv A\left[\frac{A\phi r z}{\eta}\right]^{\frac{-\phi}{\phi+\frac{r}{\mu}-1}}\left(\frac{\phi\mu}{\mu-r}-1\right)-\frac{\eta\mu}{r(\mu-r)z};$$
$$\widehat{B}(z) \equiv \frac{\mu}{r(\mu-r)}\left(\frac{\eta}{z}\right)^{\frac{r}{\mu}}-A\left(\frac{z}{\eta}\right)^{\phi}-\frac{\eta\mu}{r(\mu-r)z}.$$

Proposition 1 *Suppose $N=1$. When (12) is satisfied and $B<\widetilde{B}(\delta_0)$, an optimal downward barrier adjustment will be made at $\widehat{t}$ and $\delta_1^*=\theta(\delta_0)\delta_0$, where the reciprocal of the adjustment size is given by*

$$\theta(\delta_0) \equiv \left[\frac{A\phi r\delta_0}{\eta}\right]^{\frac{1}{\phi+\frac{r}{\mu}-1}}. \tag{13}$$

When $A\phi r \leq (A\phi r)^{\frac{1}{\phi+\frac{r}{\mu}}} < \frac{\eta}{\delta_0}$ and $B<\widehat{B}(\delta_0)$ hold, an optimal downward barrier adjustment will be made at $\widehat{t}$ and $\delta_1^*=\eta$. Otherwise no adjustment will be made. Correspondingly, the value function is given by

$$V_1(\delta_0,h_0)=\begin{cases} \frac{\mu\eta}{r(\mu-r)}\left(\frac{\eta}{\delta_0 h_0}\right)^{\frac{-r}{\mu}}\theta(\delta_0)^{\frac{r}{\mu}-1}\delta_0^{-1} & \\ -\frac{h_0}{\mu-r}-\left(\frac{\eta}{\delta_0 h_0}\right)^{\frac{-r}{\mu}}(A\theta(\delta_0)^{-\phi}+B), & \text{when} \quad \widetilde{B}(\delta_0)>B \text{ and (12) is satisfied.} \\ \frac{\mu}{r(\mu-r)}(h_0)^{\frac{r}{\mu}}-\frac{h_0}{\mu-r} & \\ -\left(\frac{\eta}{\delta_0 h_0}\right)^{\frac{-r}{\mu}}(A(\frac{\eta}{\delta_0})^{-\phi}+B), & \text{when} \quad \widehat{B}(\delta_0)>B \text{ and } A\phi r\leq(A\phi r)^{\frac{1}{\phi+\frac{r}{\mu}}}<\frac{\eta}{\delta_0} \\ V_0(\delta_0,h_0), & \text{otherwise} \end{cases}. \tag{14}$$

Proof. Refer to Appendix 3. ∎

The intuition is easy to understand by checking the option value of having one adjustment opportunity. Suppose the current institutional variable is δ and the learning constraint is already binding (i.e., $x=\frac{\eta}{\delta}$). The instantaneous net gain by adjusting δ to a value $\widetilde{\delta}\in[\eta,\delta)$ is given by $V_0(\widetilde{\delta},\frac{\eta}{\delta})-C(\delta,\widetilde{\delta})-V_0(\delta,\frac{\eta}{\delta})$. Let $y\equiv\frac{\delta}{\widetilde{\delta}}$ denote the adjustment size and $\Omega(y,\delta)$ denote the current value of the net gain by undertaking an adjustment with size y from δ, which is given by

$$\Omega(y,\delta) \equiv \frac{\mu\eta}{\delta r(\mu - r)}\left[y^{1-\frac{r}{\mu}} - 1\right] - \left[Ay^{\phi} + B\right]. \qquad (15)$$

An adjustment will be exercised if and only if there exists some $\widehat{y} \in (1, \frac{\delta}{\eta}]$ such that $\Omega(\widehat{y},\delta) > 0$. The option value of having one adjustment opportunity is $\max\{0, \max_{y\in(1,\frac{\delta}{\eta}]} \Omega(y,\delta)\}$. Let $\widetilde{y}$ denote the smallest positive root of $\Omega(y,\delta) = 0$ for any given δ if there exists some $\widehat{y} \in (1, \frac{\delta}{\eta}]$ such that $\Omega(\widehat{y},\delta) > 0$. Observe that $\lim_{y\downarrow 1}\Omega(y,\delta) = -(A+B) < 0$ and $\Omega(y,\delta)$ is strictly increasing in y on $(1, \theta(\delta)^{-1})$. So when $\widetilde{B}(\delta) > B$, the Mean-Value Theorem implies that there exists a unique root of $\Omega(y,\delta)$, denoted as $\widetilde{y}(\delta)$.

Now imagine there are only N adjustment opportunities, so after the first $N-1$ opportunities have been used, the social planner is left with only one option to make adjustment. Figure 2 illustrates the optimal decision for this last option of reform.

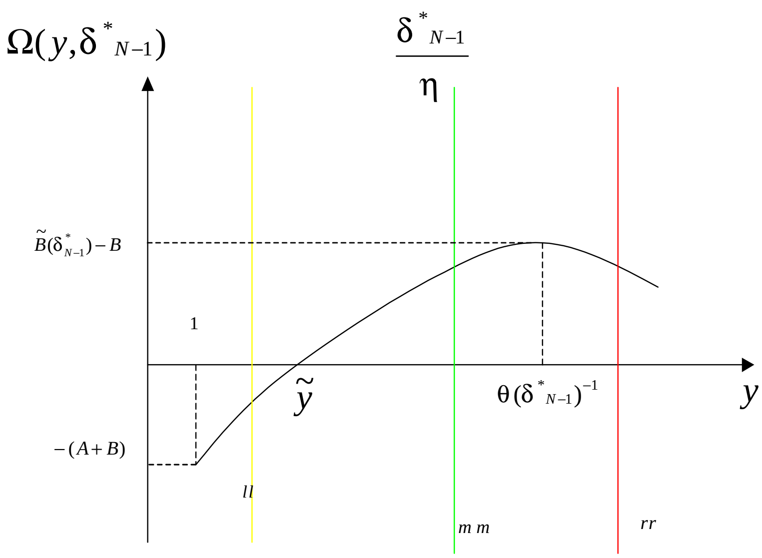

Figure 2. Last Adjustment Option

Suppose δ^*_{N-1} is small enough such that $\frac{\delta^*_{N-1}}{\eta} < \widetilde{y}(\delta^*_{N-1})$. That is, the vertical line $y = \frac{\delta^*_{N-1}}{\eta}$ is in a position like ll in Figure 2. Then the net benefit from any adjustment must be no larger than $\Omega(\frac{\delta^*_{N-1}}{\eta}, \delta^*_{N-1})$, which is negative. This implies that it is optimal to waive the last option of adjustment. In that case, the long run steady state of the institutional variable is δ^*_{N-1}. If $\widetilde{y}(\delta^*_{N-1}) < \frac{\delta^*_{N-1}}{\eta} \leq \theta(\delta^*_{N-1})^{-1}$ is true so that $y = \frac{\delta^*_{N-1}}{\eta}$ is in a position like mm, then the last adjustment is made, after which $\delta^*_N = \eta$ and the developing country will eventually have the same GDP per capita as the developed country. If $\frac{\delta^*_{N-1}}{\eta} > \theta(\delta^*_{N-1})^{-1}$ is true so that $y = \frac{\delta^*_{N-1}}{\eta}$ is in a position like rr, then the last adjustment opportunity is also used and the long run barrier variable will become $\delta^*_N = \theta(\delta^*_{N-1})\delta^*_{N-1}$, which is larger than η. In this case, a permanent GDP gap exists between the two countries. This is because the barrier adjustment becomes increasingly costly while the benefit for the same adjustment size decreases as the economy grows. Eventually the GDP gap becomes sufficiently small that any further adjustment becomes unattractive.

In particular, when $N = 1$, the first inequality in (12) ensures $\theta(\delta_0)^{-1}$ is greater than one, while the second weak inequality guarantees that $\theta(\delta_0)^{-1}$ is to the left of the vertical line $y = \frac{\delta_0}{\eta}$ (in a position like rr). $B < \widetilde{B}\ (\delta_0)$ guarantees that the option value of having one adjustment opportunity is strictly positive. When the second weak inequality in (12) is violated, there are two possibilities. One is $\widehat{B}(\delta_0) > B$, in which case line $y = \frac{\delta_0}{\eta}$ is at a position like mm (that is, $\frac{\delta_0}{\eta} > \widetilde{y}(\delta_0)$) so the barrier variable is adjusted to η. The other possibility is when $\widehat{B}(\delta_0) \leq B$, in which case line $y = \frac{\delta_0}{\eta}$ is at a position like ll so the adjustment option is waived. For all the remaining circumstances, no adjustment is made. This completes the geometric illustration for $N = 1$, as characterized in (14).

The conditions in (14) are complicated. Alternatively, the following lemma gives a useful and easy-to-check necessary condition to exercise the one-shot control.

Lemma 3. *The one-time control will be exercised only if*

$$A + B < \frac{1}{r}\left[\frac{\mu}{r}\right]^{\frac{r}{r-\mu}}, \tag{16}$$

and $\frac{\eta}{\delta_0} \in (\ \underline{\beta}, \overline{\beta})$ *, where* $\underline{\beta}$ *and* $\overline{\beta}$ *are the two distinct roots of the following equation:*

$$x^{\frac{r}{\mu}} = x + \frac{(A+B)\, r(\mu - r)}{\mu}.$$

Proof. Refer to Appendix 4. ■

To understand (16), observe that any adjustment at least costs $A + B$, so it has to be sufficiently small to warrant a reform. Moreover, if $\frac{\eta}{\delta_0} \geq \overline{\beta}$, then δ_0 is sufficiently close to η that the benefit from any further adjustment is too small to warrant further adjustment. If $\frac{\eta}{\delta_0} \leq \underline{\beta}$, then no further one-step adjustment will be made because δ_0 is so high that it requires a large reduction in δ to achieve any given amount of utility improvement, making the cost of the associated adjustment larger than the gain from any one-step adjustment.

4.3 Optimal Reform

The value function with N adjustment options, where $1 \leq N < \infty$, is given by

$$V_N(\delta_0, h_0) = \max\{G_N(\delta_0, h_0), F_N(\delta_0, h_0)\},$$

where

$$G_N(\delta_0, h_0) \equiv \max_{t_1 \leq \widehat{t}, \delta_1} \int_0^{t_1} h_0 e^{\mu t} e^{-rt} dt + e^{-rt_1} \left[V_{N-1}(\delta_1, h_0 e^{\mu t_1}) - C(\delta_0, \delta_1) \right];$$

$$F_N(\delta_0, h_0) \equiv \max_{\widehat{t} \leq t_1, \delta_1} e^{-rt_1} \left[V_{N-1}(\delta_1, \frac{\eta}{\delta_0}) - C(\delta_0, \delta_1) - V_0(\delta_0, \frac{\eta}{\delta_0}) \right] + V_0(\delta_0, h_0).$$

I obtain the following result by using the recursive method:

Proposition 2 *Any institutional adjustment must occur precisely when the learning constraint just becomes binding, that is,*

$$t_i^* = \frac{1}{\mu} \ln \frac{\eta}{\delta_{i-1}^* h_0}, \forall i = 1, 2, ..., N. \tag{17}$$

In addition, the developing economy keeps catching up at a constant speed μ until the last learning constraint becomes binds, after which convergence stops, that is

$$\frac{\dot{h}(t)}{h(t)} = \begin{cases} \mu, & when \quad t \leq t_N^* \\ 0, & otherwise \end{cases}.$$

Proof. Refer to Appendix 5. ■

This proposition states that economic convergence is accompanied by a process of cumulative institutional buildings. The institutional barrier is sequentially reduced in a timely manner to ensure a generically unbinding learning constraint. In equilibrium the GDP dynamics also appear to be solely determined by the human capital accumulation, as argued in standard endogenous growth literature. However, what this model highlights is the crucial and hidden role played by the cumulative institutional building in sustaining this convergence process. Without the timely relaxations of institutional binding constraints at different development level, convergence will stop prematurely. This fundamental interaction between institutional reforms and economic growth seems largely ignored in the existing convergence literature.

Recall that the "Washington Consensus" emphasizes that all the reforms should be undertaken in one step so that all the future growth will be free of any binding institutional bottlenecks (Stiglitz (1998)). It is also often argued that gradual and partial reforms may creative more distortions, so reforms should be comprehensive and quick (Murphy et al (1992)). By contrast, the model developed here formalizes a rationale for why optimal reforms can be done sequentially along with the convergence. The model predictions for reform and convergence are quite consistent with the Chinese experience discussed in Section 2 as well as many cross-country real-life episodes of accelerations and reforms (Sachs and Warner (1995), Rodrik (2005), Wade (2000), World Bank (2005)).

Moreover, if a binding learning constraint can be interpreted as a "crisis" since the constraint can potentially strangle further convergence, then Proposition 2 is also congruent with the "crisis hypothesis" empirically established in the reform literature, which states that the reform is more likely to occur when a "crisis" appears (Drazen and Grilli (1993), Alesina et al (2006)).

In the model, cumulative reforms are needed to sustain the convergence, but in equilibrium do reforms occur infinite times? To address this issue, I define $N^* \equiv \inf\left\{\arg\max_N V_N(\delta_0, h_0)\right\}$, the optimal minimum number of institutional adjustments.

Proposition 3 *There is only a finite number of reforms ($N^* < \infty$).*

Proof. Refer to Appendix 6. ■

To understand the intuition for why the Ramsey government chooses to only conduct a finite number of reforms, first note that the total potential gain of reform is finite. In addition, (17) in the previous proposition implies that no reform would occur after $\frac{-\ln h_0}{\mu}$, which means that each desirable reform must entail a minimum positive cost with the present discounted value strictly larger than $(A+B)\, e^{\frac{-\ln h_0}{\mu}}$, so it does not pay to do an infinite number of reforms. Notice that this is true as long as A and B are not both zero simultaneously. It implies that convergence stops at some finite time point. Methodologically, this proposition also warrants the method of backward induction employed in my characterization.

Suppose $N^* \geq 1$. The original optimization problem (5) can be rewritten as

$$\max_{N, \{\delta_i\}_{i=1}^N} V_0(\delta_N, h_0) - \sum_{i=1}^{N} e^{-rt_i} \left[A \left(\frac{\delta_{i-1}}{\delta_i} \right)^{\phi} + B \right] \tag{18}$$

subject to

$$\delta_i < \delta_{i-1} \text{ for each } i = 1, ..., N,$$
$$\delta_N \geq \eta; \;\; \delta_0 \text{ and } h_0 \text{ are given,}$$

where t_i is given by (17) for any $i = 1, ..., N$.

Recall that the optimal adjustment plan automatically satisfies the budge constraint when the international credit market is complete. Substituting (17) into (18) and using (8) yield the following equivalent problem:

$$\max_{N, \{\delta_i\}_{i=1}^N} \frac{\mu\eta}{r(\mu - r)} \left(\delta_N\right)^{\frac{r}{\mu}-1} - \sum_{i=1}^{N} \delta_{i-1}^{\frac{r}{\mu}} \left[A \left(\frac{\delta_{i-1}}{\delta_i} \right)^{\phi} + B \right]. \tag{19}$$

Observe that N^* and $\{\delta_i^*\}_{i=1}^{N^*}$ are independent of h_0 as long as assumption A0 is satisfied. This is because, for any given δ_0, no matter what h_0 is, the economy will have the same GDP at time $\widehat{t}$. From that point on, the optimization problem is identical and independent of h_0, so the initial institutional barrier δ_0 alone will determine the optimal adjustment scheme. When $A = 0$, (19) implies that any reform, if initiated

($N^* > 0$), should be undertaken thoroughly once and for all ($N^* = 1$ and $\delta_N^* = \eta$). By revoking (14), I obtain that $N^* = 1$ if and only if $\widehat{B}(\delta_0) > B$, or equivalently, $\frac{\mu}{r(\mu-r)}\left[\left(\frac{\eta}{\delta_0}\right)^{\frac{r}{\mu}} - \frac{\eta}{\delta_0}\right] > B$. Otherwise, no reform occurs. For exposition convenience, let $y_i \equiv \frac{\delta_{i-1}}{\delta_i}$ denote the size of the ith institutional adjustment for any positive integer $i \leq N^*$ and define $\mathcal{N} \equiv \{1, 2, ..., N^* - 1\}$.

Proposition 4 *Suppose $A > 0$ and $N^* \geq 2$.*[16] *The reform size changes monotonically over time. More precisely, when $\delta_{N^*}^* > \eta$, the following is true:*

$$\frac{y_{i+1}}{y_i} \gtreqless 1, \forall i \in \mathcal{N}, \text{ iff } y_j \gtreqless \varpi, \text{ for some } j \in \mathcal{N}, \tag{20}$$

where ϖ is uniquely determined by

$$\omega^{\frac{r}{\mu}}\phi - \frac{B}{A}\frac{r}{\mu\varpi^{\phi}} = \frac{r}{\mu} + \phi \tag{21}$$

When $\delta_{N^*}^* = \eta$, the adjustment is also strictly monotonic or constant over time.

Proof. Refer to Appendix 8. ■

This proposition states that, whenever multiple reforms are conducted, the magnitudes of the reforms either increase or diminish over time. Based on (2), the monotonicity of the reform sizes implies that reforms are more frequent when the reform sizes diminish over time because the convergence period supported by each reform becomes shorter and shorter. Likewise, if the reform sizes increase over time, then the frequency of reforms becomes smaller.

Another immediate implication of this proposition is that either the reform sizes are all above ϖ or all below ϖ, the value of which is determined by (21). Moreover, if a GDP gap exists in the long run ($\delta_{N^*}^* > \eta$), then $y_N = \theta^{-1}(\delta_{N-1})$ and $\delta_N = \delta_{N-1}y_N^{-1}$. Based on Proposition 4, it can be shown that the long-run GDP per capita is larger than $A\phi r\varpi^{\phi+\frac{r}{\mu}}$ if and only if the reform sizes are monotonically increasing, and the GDP per capita is smaller than $A\phi r\varpi^{\phi+\frac{r}{\mu}}$ if the reform sizes are monotonically decreasing. The GDP per capita equals $A\phi r\varpi^{\phi+\frac{r}{\mu}}$ in the long run if and only if the reform sizes are constant, in which case

$$\delta_N = \frac{\eta}{A\phi r}\varpi^{-\left(\phi+\frac{r}{\mu}\right)} = \delta_0\varpi^{-N}$$

therefore

$$N^* = \frac{\log\frac{\delta_0 A\phi r}{\eta}}{\log\varpi} + \phi + \frac{r}{\mu}$$

It implies that $\frac{\partial N^*}{\partial\delta_0} > 0$, so the higher the initial institutional barrier, the more reforms there will be; $\frac{\partial N^*}{\partial\eta} < 0$, implying that more efficient absorption of learning externality

[16] A full characterization for the case with $N = 2$ is provided in the Appendix 7.

reduces the number of reforms. Moreover, $\frac{\partial N^*}{\partial \phi} > 0$ and $\frac{\partial N^*}{\partial A} > 0$, indicating that the less distorting the reform process (smaller ϕ or A), the fewer steps of reforms are needed. $\frac{\partial N^*}{\partial B} < 0$, so a larger fixed cost leads to fewer reforms. These results is intuitive: when the variable adjustment cost becomes relatively important, it is better to reduce the size of adjustment size (and also make the reforms more frequent). So if the less developed economy has a very powerful single-party administrative central government (think about China), then it has a relatively small B but a relatively big A (also see Besley and Kudamatsu (2008)). The model implies that the optimal reform for this economy should be more piece-meal. Even within democracies, a proportional representation parliamentary system can be very different from a presidential system. The latter tends to have a smaller B than a parliament system and therefore the model predicts that a democracy with presidential system should adopt a more gradual small-step reform than the countries with proportional representation system, holding everything else equal.

Dixit (2004) argues that setting up formal institutions (such as legal rules and democratic political systems) requires a large B but a relatively small A, whereas the opposite is true for informal institutions. With this interpretation, the model implies that the optimal reform tends to be quicker when many informal institutions are still on the reform list, especially at the early stage of reform, but the reform gets slower when formal institutions need changing, presumably at the later stage of reform. All these predictions are testable empirically.

5 Reform Under Incomplete Credit Market

Perfect international credit market allows us to essentially ignore the budget constraint problem. However, when the international credit market is imperfect in the sense that the developing country cannot borrow in the international market, then the country has to rely on its own saving to finance its institutional adjustment. Since the Ramsey government can freely postpone consumption because $\rho = r$, consumption can be always zero before the final reform, just to avoid the binding budget constraint problem as much as possible. That is, $c(s) = 0, \forall s < T^{**}$, where T^{**} demotes the time of the last adjustment. The feasibility constraint (6) can be rewritten as follows

$$\int_0^{\widetilde{t}_{i+1}} h(t)e^{-rt}dt - \sum_{j=0}^{i} C(\delta_j, \delta_{j+1})e^{-r\widetilde{t}_{j+1}} \geq 0 \text{ for each } i = 0, 1, 2, ..., \tag{22}$$

where $\widetilde{t}_i$ denotes the time point of the i-th adjustment under the imperfect credit market. Let $\widetilde{N}$ denote the minimum optimal number of adjustments. Obviously, $\widetilde{N} < \infty$ because the same logic in the proof of Proposition 3 remains valid.

The Ramsey government now maximizes the total present discounted value of consumption (1) subject to (2), (3), with δ_0 and h_0 given, and subject to a sequence of budget constraints given by (22).

Observe that T^{**}(or equivalently $\widetilde{t}_{\widetilde{N}}$) must be finite, otherwise the net benefit of the last reform is not strictly positive, which contradicts (22). If the optimal adjustment scheme obtained in the last section (with perfect international credit market) automatically satisfies (22), then that scheme and the associated growth dynamics will be also the optimal ones under the imperfect credit market. Otherwise, the constrained optimization must be newly analyzed.

I consider the simplest case, in which there is only one opportunity to reform.[17] The value function becomes

$$\widetilde{V}_1(\delta_0, h_0) = \max\{\widetilde{G}(\delta_0, h_0), \widetilde{F}(\delta_0, h_0)\}, \tag{23}$$

where

$$\widetilde{G}(\delta_0, h_0) \equiv \max_{\widetilde{t}_1 \leq \widehat{t}, \delta_1} \int_0^{\widetilde{t}_1} h_0 e^{\mu t} e^{-rt} dt + e^{-r\widetilde{t}_1} \left[V_0(\delta_1, h_0 e^{\mu \widetilde{t}_1}) - C(\delta_0, \delta_1) \right] \tag{24}$$

subject to

$$\int_0^{\widetilde{t}_1} h_0 e^{\mu t} e^{-rt} dt \geq e^{-r\widetilde{t}_1} C(\delta_0, \delta_1) \tag{25}$$

and

$$\widetilde{F}(\delta_0, h_0) \equiv \max_{\widehat{t} \leq \widetilde{t}_1, \delta_1} \int_0^{\widehat{t}} h_0 e^{\mu t} e^{-rt} dt + \int_{\widehat{t}}^{\widetilde{t}_1} \frac{\eta}{\delta_0} e^{-rt} dt + e^{-r\widetilde{t}_1} \left[V_0(\delta_1, \frac{\eta}{\delta_0}) - C(\delta_0, \delta_1) \right] \tag{26}$$

subject to

$$\int_0^{\widehat{t}} h_0 e^{\mu t} e^{-rt} dt + \int_{\widehat{t}}^{\widetilde{t}_1} \frac{\eta}{\delta_0} e^{-rt} dt \geq e^{-r\widetilde{t}_1} C(\delta_0, \delta_1), \tag{27}$$

and as before, I will focus on the case when $\eta > x_0 \delta_0$.

The same logic in Lemma 1 still applies under the imperfect credit market, so any adjustment ($0 < \widetilde{t}_1 < \infty$) must be made weakly after the learning constraint is binding: $\widetilde{V}_1(\delta_0, h_0) = \widetilde{F}(\delta_0, h_0)$. (26) can be rewritten as

$$\widetilde{F}(\delta_0, h_0) = \max_{\widehat{t} \leq \widetilde{t}_1, \delta_1} V_0(\delta_0, h_0) + e^{-r\widetilde{t}_1} \left[V_0(\delta_1, \frac{\eta}{\delta_0}) - C(\delta_0, \delta_1) - V_0(\delta_0, \frac{\eta}{\delta_0}) \right]. \tag{28}$$

Proposition 5 *Under the imperfect credit market, when (12) and $B < \widetilde{B}\ (\delta_0)$ are true, the first-best reform size (given by (13)) is implemented at time $\widehat{t}$ if and only if the following two conditions are true:*

$$\frac{\eta}{(\mu - r)\delta_0} - A \left[\frac{A\phi r \delta_0}{\eta} \right]^{\frac{-\phi}{\phi + \frac{r}{\mu} - 1}} > B, \tag{29}$$

[17] Characterization of multiple reforms under imperfect credit market is much more complicated and hence reserved for future research.

and $h_0 \leq h^*$, where h^* is given by

$$h^* \equiv \left\{ \left(\frac{\eta}{\delta_0}\right)^{\frac{-r}{\mu}} \left[\frac{\eta}{\delta_0} - (\mu - r)(A\left[\frac{A\phi r\delta_0}{\eta}\right]^{\frac{-\phi}{\phi+\frac{r}{\mu}-1}} + B)\right]\right\}^{\frac{\mu}{\mu-r}}.$$

Proof. Refer to Appendix 9.Under assumption (12) and $B < \widetilde{B}\ (\delta_0)$, the optimal adjustment size given by (13) can be fully financed by the domestic saving of the developing economy if and only if $Q(h_0) \geqslant 0$, where

$$Q(z) \equiv \frac{z}{(\mu - r)}\left[\frac{\eta}{\delta_0 z} - \left(\frac{\eta}{\delta_0 z}\right)^{\frac{r}{\mu}}\right] - (A\left[\frac{A\phi r\delta_0}{\eta}\right]^{\frac{-\phi}{\phi+\frac{r}{\mu}-1}} + B).$$

Obviously, $Q(z) < 0$ for any z when (29) is violated. In that case, the first-best reform is never feasible. Observe that $Q'(z) < 0$ whenever $z > 0$. There is a unique root, h^*,of the equation $Q(h^*) = 0$. It can be verified that $h^* < \frac{\eta}{\delta_0}$ because of (29). Thus $Q(h_0) > 0$ iff $h_0 < h^*$. ■

This proposition says that whether self financing is enough shall depend on the initial human capital stock h_0. When the initial endowment of human capital is sufficiently large, then the first-best reform size cannot be sufficiently financed. This appears counter-intuitive as one may think that higher initial endowment would imply a higher financing ability. This paradox can be resolved as follows. Although the optimal size of reform is independent of the initial human capital under perfect credit market (recall (13)), the optimal timing of reform does depend on the initial factor endowment (wealth). (7) suggests that the higher the initial human capital, the sooner the learning constraint becomes binding, so the present discounted cost of the first-best reform size becomes higher. It turns out that this timing effect dominates the initial endowment effect, therefore, under imperfect credit market, the learning constraint binds before enough saving is accumulated to cover the cost for the first-best reform size.

Now if $h_0 > h^*$, how does the developing economy modify its plan under the financial constraint? Solving (28) and (27) by using the Lagrangian, I obtain that, if an adjustment is made and (27) binds, the optimal barrier target δ_1 is determined by the following equation:

$$\frac{\frac{\partial V_0(\delta_1,h_0)}{\partial\delta_1}}{V_0(\delta_1, h_0)} = \frac{\frac{\partial[C(\delta_0,\delta_1)+V_0(\delta_0,\frac{\eta}{\delta_0})]}{\partial\delta_1}}{C(\delta_0, \delta_1) + V_0(\delta_0, \frac{\eta}{\delta_0})}, \tag{30}$$

which states that the marginal percentage increase in the value due to the barrier adjustment is equal to the marginal percentage increase in the total opportunity cost, which comprises the direct adjustment cost, $C(\delta_0, \delta_1)$, and the foregone utility level without institutional adjustment, $V_0(\delta_0, \frac{\eta}{\delta_0})$. The main reason why we have the equality between the percentage changes in benefit and opportunity cost is due to that the budget constraint (30) binds. Any beneficial adjustment implies $V_0(\delta_1, h_0) > C(\delta_0, \delta_1) + V_0(\delta_0, \frac{\eta}{\delta_0})$, so the optimal δ_1 must satisfy $\frac{\partial V_0(\delta_1,h_0)}{\partial\delta_1} > \frac{\partial C(\delta_0,\delta_1)}{\partial\delta_1}$, confirming that $\delta_1 > \delta_0\theta(\delta_0)$. (30) can be rewritten as

$$M\delta_1^{\frac{r}{\mu}+\phi-1} + H\delta_1^{\frac{r}{\mu}-1} + \frac{A\phi\delta_0^{\phi}h_0}{\mu-r} = 0, \tag{31}$$

where

$$M \equiv (B + \frac{\eta}{r\delta_0})\frac{\eta}{r}\frac{\eta}{h_0},\ \ H \equiv A\delta_0^{\phi}\frac{\eta}{r}\left(\frac{\eta}{h_0}\right)^{-\frac{r}{\mu}}(1-\frac{\phi\mu}{\mu-r}), \delta_1 \in [\eta, \delta_0).$$

The left hand side of equation (31) is strictly increasing in δ_1, so there is at most one solution.

The previous analysis has already shown that the option value of having one opportunity to reform is positive and equal to $\Omega(\theta(\delta_0)^{-1}, \delta_0)$ when $B < \widetilde{B}\ (\delta_0)$. Thus an

adjustment has to be made because any adjustment size would become feasible in the long run. When the solution to (31) exists and satisfies $\delta_1^* \in [\eta, \delta_0)$, it must also satisfy $\Omega(\frac{\delta_0}{\delta_1^*}, \delta_0) > 0$. The optimal adjustment time is determined by the binding budget constraint and given by

$$\widehat{t}_1^* = \frac{1}{r}\ln\frac{C(\delta_0, \delta_1^*) + V_0(\delta_0, \frac{\eta}{\delta_0})}{V_0(\delta_0, h_0)}. \tag{32}$$

The above findings are summarized as follows:

Proposition 6 *Suppose $\eta/\delta_0 > h_0 > h^*$. Suppose the international credit market is incomplete and reform can be implemented at most once. Under assumptions (12), (29), (40) and $B < \widetilde{B}\ (\delta_0)$, a reform will be made at time $\widehat{t}_1^*$ given by (32) and δ_1^* is uniquely determined by (31).*

To illustrate this proposition more intuitively, I need to develop more notations first. Let $\widetilde{\delta}_1$ denote the largest adjustment target affordable at the first binding time point $\widehat{t}$. When the following is true:

$$\frac{h_0}{(\mu-r)}\left[\frac{\eta}{\delta_0 h_0} - \left(\frac{\eta}{\delta_0 h_0}\right)^{\frac{r}{\mu}}\right] > A + B, \tag{33}$$

the binding budget constraint implies $\widetilde{\delta}_1 = \delta_0\psi(\delta_0)$, where

$$\psi(\delta_0) \equiv \left[\frac{\frac{h_0}{(\mu-r)}\left[\frac{\eta}{\delta_0 h_0} - \left(\frac{\eta}{\delta_0 h_0}\right)^{\frac{r}{\mu}}\right] - B}{A}\right]^{-\frac{1}{\phi}}.$$

This proposition can be illustrated by Figure 3. $h_0 > h^*$ makes sure that $\psi(\delta_0)^{-1} <$

$\theta(\delta_0)^{-1}$.(33) ensures that $\psi(\delta_0)$ is well defined and $\psi(\delta_0)^{-1} > 1$. At time $\widehat{t}$, the set of the affordable adjustment sizes is $[1, \psi(\delta_0)^{-1}]$. When $\psi(\delta_0)^{-1} < \widetilde{y}(\delta_0)$ ((that is, line

$y = \psi(\delta_0)^{-1}$ is at a position like line lll), $\Omega(y, \delta_0) < 0$ for any $y \in [1, \psi(\delta_0)^{-1}]$, thus no affordable adjustment is profitable enough to compensate the adjustment cost at $\widehat{t}$. Nevertheless, $B < \widetilde{B}\ (\delta_0)$ ensures that the option value of the adjustment opportunity is strictly positive, so the one-time reform must be made because any adjustment size becomes affordable when savings last for a sufficiently long period. Thus the reform must occur strictly after the learning constraint remains binding for a while.

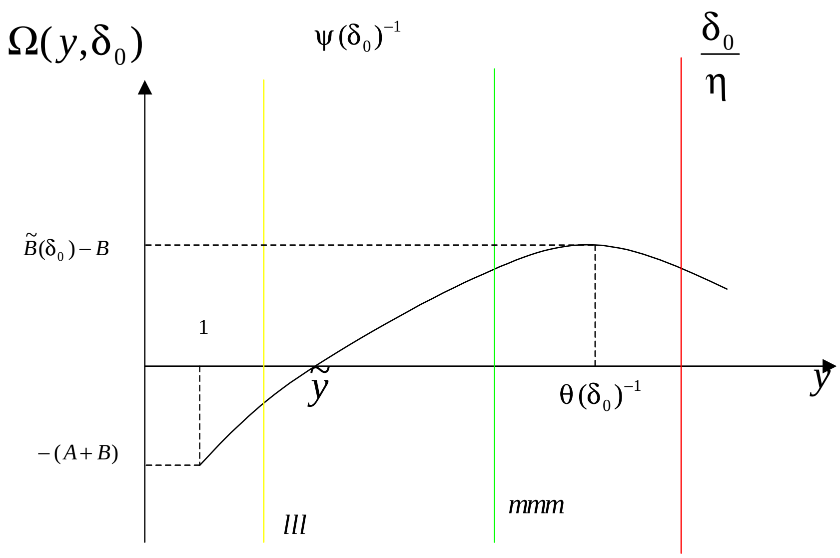

Figure 3. Optimal Reform under Incomplete Credit Market

If $\psi(\delta_0)^{-1} > \widetilde{y}(\delta_0)$ (that is, line $y = \psi(\delta_0)^{-1}$ is at a position like line mmm), then a profitable adjustment is affordable at $\widehat{t}$. The monotonicity of $\Omega(\cdot, \delta_0)$ on the relevant interval implies that if an adjustment is made at $\widehat{t}$, it must be adjusted to $\psi(\delta_0)\delta_0$, the largest feasible adjustment size.

To better characterize the optimal reform time, let $\widetilde{T}_1^*$ denote the time at which the first-best adjustment size (13) first becomes affordable, thus $\widetilde{T}_1^*$ is determined by

$$\int_0^{\widehat{t}} h_0 e^{\mu t} e^{-rt} dt + \int_{\widehat{t}}^{\widetilde{T}_1^*} \frac{\eta}{\delta_0} e^{-rt} dt = e^{-r\widetilde{T}_1^*} C(\delta_0, \theta(\delta_0)\delta_0). \tag{34}$$

No reform is later than $\widetilde{T}_1^*$ because the first-best adjustment size (13) is affordable by $\widetilde{T}_1^*$ and any further delay is costly. In other words, $\widehat{t}_1^* \in [\widehat{t}, \widetilde{T}_1^*]$. In addition, (32) shows that, when the reform has a larger size (smaller δ_1^*), it occurs later (larger $\widehat{t}_1^*$). Notice that (31) implies

$$\frac{\partial \delta_1^*}{\partial A} > 0; \frac{\partial \delta_1^*}{\partial B} < 0; \frac{\partial \delta_1^*}{\partial \delta_0} > 0; \frac{\partial \delta_1^*}{\partial h_0} > 0.$$

Clearly, $\widehat{t}_1^* > \widehat{t}$ when $\psi(\delta_0)^{-1} < \widetilde{y}(\delta_0)$, which implies that the developing economy stops converging at time $\widehat{t}$ and remains so until the reform occurs at $\widehat{t}_1^*$, after which convergence resumes. This pattern of punctuated convergence is different from the continuous convergence under the perfect international credit market, or when sufficient foreign support is available.[18]

6 Concluding Remarks

There is compelling cross-country empirical evidence that economic accelerations in developing economies often kick start after a modest policy or institutional adjustment rather than a fundamental reform across the board. It is also observed that sustained convergence of those successful chasers are featured by a process of successive reforms that relax different and sequentially binding growth bottlenecks as the economy develops. Reforms support convergence, which in turn leads to new binding constraints and thus triggers new reforms. It is economic development that turns some previously relaxed constraint into a binding bottleneck. Then it requires a further reform to eliminate the barrier to sustain the catch-up growth. China is a case in point. I develop a stylized model within the growth framework to formalize this interactive process between convergence and reforms. Using recursive methods, I characterize this technically non-trivial dynamic reform problem faced by an artificial benevolent social planner and I also fully characterize the resulting growth pattern.

In this normative investigation, the predictions and prerequisites of the first-best reform are all explicitly specified and organized in a logically coherent way, which facilitates future explorations. Qualitatively, it seems fruitful to introduce uncertainty into the model, which is indispensable if we want to capture the experimental and pragmatic nature of China's reform and growth more accurately (Hausmann et al (2008), Rodrik (2010)) or if macro volatility and risk tolerance are major concerns (Fernandez and Rodrik (1991), Dewatripont and Roland (1995)). Another promising avenue is to introduce more explicit political economy elements such as conflicting groups or selfish leaders/reformers (Wei (1997), Acemoglu (2005), Roland (2000), Acemoglu (2005), Li et al (2012)), or to explicitly discuss both economic and political liberalizations (Caselli and Gennaioli (2008), Giavazzi and Tabellini (2005)). On the quantitative side, empirical investigations are called for to assess the performance of the current model or its extensions.

[18] However, Easterly (2005) argues that in reality loans and foreign aid in general do not help structural adjustment and economic growth in most recipient countries.

References

[1] Acemoglu, Daron, 2005. "Politics and Economics in Weak and Strong States”, *Journal of Monetary Economics* 52: 1199-1226

[2] ——, Simon Johnson, and James A. Robinson, 2005. “Institutions as the Fundamental Causes of Long-run Growth”, *Handbook of Economic Growth* (Philippe Aghion and Stephen Durlauf, eds., North Holland)

[3] ——, Philippe Aghion, and Fabrizio Zilibotti. 2006. "Distance to Frontier, Selection, and Economic Growth", *Journal of the European Economic Association* 4: 37-74.

[4] Alesina, Alberto, and Allan Drazen.1991. “Why are Stabilizations Delayed?” *American Economic Review* 81(5):1170-81.

[5] Alesina, Alberto, Silvia Ardagna, and Francesco Trebbi, 2006. " Who Adjusts And When? On the Political Economy of Stablizations" , *IMF Staff Papers,* Mundell-Flemming Lecture 53: 1-49

[6] Banerjee, Abhijit, Maitreesh Ghatak, and Mark Gertler, 2002. “Empowerment and Efficiency: Tanancy reform in West Bengal,” *Journal of Political Economy*, 110(2): 239-80.

[7] Barberis, Nicholas, Maxim Boycko, Andrei Shleifer and Natalia Tsukanova, 1996. “How Does PrivatizationWork? Evidence from The Russian Shops,” *Journal of Political Economy* 104(4): 764-790.

[8] Barro, Robert J. and Xavier Sala-I-Martin, 1992. "Convergence", *Journal of Political Economy* 100(2): 223-251

[9] Beck, Torsten, and Luc Laeven, 2006. “Institution Building and Growth in Transition Economies,” *Journal of Economic Growth* 11: 157–186

[10] Benhabib, Jess and Mark M Spiegel, 2005. "Human Capital and Technology Diffusion", *Handbook of Economic Growth,* Philippe Aghion and Stephen Durlauf, eds., North Holland.

[11] Besley, Timothy and Masa Kudamatsu. 2008. "Making Autocracy Work," in *Institutions and Economic Performance*, edited by Elhanan Helpman, Harvard University Press.

[12] Besley, Timothy and Torsten Persson. 2009. "The Origins of State Capacity: Property Rights, Taxation, and Policy", *American Economic Review* 99, 1218-1244, 2009

[13] Blanchard, Olivier and Andrei Shleifer. 2001. "Political Centralization and Fiscal Federalism", working paper, MIT

[14] Canda, Vandana. 2006. *Technology, Adaptation, and Exports: How Some Developing Countries Got It Right.* The World Bank Publications

[15] Caselli, Francesco and Nicola Gennaioli, 2008. "Economics and Politics of Alternative Institutional Reforms," *Quarterly Journal of Economics,* 123 (3): 1197-1250

[16] Dewatripont, M. and Gerard Roland, 1992. "Economic Reform and Dynamic Political Constraints" , *Review of Economic Studies* 59(4): 703-730

[17] ______. 1995. "The Design of Reform Packages under Uncertainty," *American Economic Review* 85: 1207-1223.

[18] Dixit, Avinash. 2004. *Lawlessness and Economics: Alternative Modes of Governance* . New Jersey: Princeton University Press

[19] Drazen, Allan and Vittortio Grilli, 1993. "The Benefits of Crisis for Economic Reform", *American Economic Review* 83(3): 598-607

[20] Easterly, William, 2001, "The Lost Decades: Developing Countries' Stagnation in Spite of Policy Reform 1980-1998", *Journal of Economic Growth* 6(2): 135-157

[21] ____, 2005. "What did Structural Adjustment Adjust? The Association of Policies and Growth with Repeated IMF and World Bank Adjustment Loans," *Journal of Development Economics* 76: 1-22

[22] Fernandez, Raquel and Dani Rodrik, 1991. "Resistance to Reform: Status Quo Bias in the Presence of Individual-Specific Uncertainty," *American Economic Review* 81(5): 1146-1155.

[23] Giavazzi, Francesco and Guido Tabellini, 2005. "Economic and Political Liberalizations," *Journal of Monetary Economics* 52(7):1297-1330

[24] Hall, Robert, and Charles Jones, 1999. "Why Do Some Countries Produce So Much More Output Per Worker Than Others?" *Quarterly Journal of Economics* 114:83-116.

[25] Hausmann, Ricardo, Dani Rodrik, and Andrés Velasco. 2008. "Growth Diagnostics," Chap. 15 in *The Washington Consensus Reconsidered: Towards a New Global Governance*, ed. J. Stiglitz and N. Serra. New York: Oxford University Press.

[26] Klenow, Peter J., and Andres Rodriguez-Clare, 2005. " Externality and Growth", *Handbook of Economic Growth*, Philippe Aghion and Stephen Durlauf, eds., North Holland

[27] Lau, Lawrence, Yingyi Qian, and Gerard Roland, 2000. "Reform Without Losers: An Interpretation of China's Dual-Track Approach to Transition", *Journal of Political Economy*,108(1):120-143

[28] Li, Xi, Xuewen Liu and Yong Wang. 2012. "A Model of China's State Capitalism", working paper, HKUST

[29] Lin, Justin Yifu, 1992. "Rural Reforms and Agricultural Growth in China," *American Economic Review* 82(1): 34-51

[30] ______. 2012. *Demystifying the Chinese Economy.* Cambridge: Cambridge University Press

[31] Lucas, Robert E. Jr., 2009. "Trade and the Diffusion of the Industrial Revolution", *American Economic Journal: Macroeconomics*, 1(1): 1-25

[32] Murphy, Kevin, Andrei Shleifer and Robert Vishny, 1992. " The Transition to a Market Economy: Pitfalls of Partial Reform," *Quarterly Journal of Economics* 107 (3): 889-906.

[33] Naughton, Barry. 2007. *The Chinese Economy: Transitions and Growth.* Cambridge: MIT Press

[34] Ngai, L Rachel, 2004. "Barriers and Transition to Modern Growth," *Journal of Monetary Economic*s 51:1353-1383

[35] North, Douglass. 1990. *Institutions, Institutional Change and Economic Performance,* Cambridge: Cambridge University Press

[36] Parente, Stephen L. 1994. "Technology Adoption, Learning-by-Doing, and Economic Growth", *Journal of Economic Theory*, 63(2): 346-369

[37] Parente, Stephen L and Edward C. Prescott, 1994. "Barriers to Technology Adoption and Development," *Journal of Political Economy* 102(2): 298-321.

[38] ———. 2000. *Barriers to Riches.* Cambridge: MIT Press

[39] Qian, Yingyi. 2003. "How Reform Worked in China," *In Search of Prosperity: Analytic Narratives on Economic Growth*, Dani Rodrik eds., New Jersey: Princeton University

[40] Roland, Gerard, 2000. *Transition and Economics: Politics, Markets, and Firms*, Cambridge: MIT Press

[41] Rodrik, Dani, 2005. "Growth Strategies." *Handbook of Economic Growth*, Philippe Aghion and Steven N. Durlauf, eds., North Holland: 967-1059

_____. 2010. "Diagnostics Before Prescription". *Journal of Economic Perspectives.* 24(3): 33-44.

[42] Sachs, Jeffrey D. and Andrew Warner, 1995. "Economic Reform and the Process of Global Integration", *Brookings Papers on Economic Activity*: 1-118

[43] Song, Zheng, Kjetil Storesletten, and Fabrizio Zilibotti. 2011. "Growing Like China." *American Economic Review*101(1): 196-233.

[44] Stiglitz, Joseph, 1998. "More Instruments and Broader Goals: Moving Toward the Post-Washington Consensus." World Institute for Development Economics Research Annual Lectures 2. Helsinki: WIDER

[45] Stokey, Nancy, 2012, "Catching Up and Falling Behind", working paper, University of Chicago

[46] Wade, Robert.1990. *Governing the Market: Economic Theory and the Role of Government in East Asian Industrialization.* Princeton University Press

[47] Wang, Yong, 2013. "Fiscal Decentralization, Endogenous Policies, and Foreign Direct Investment: Theory and Evidence from China and India," *Journal of Department Economics,* forthcoming

[48] Wei, Shang-Jin, 1997. "Gradualism versus Big Bang: Speed and Sustainability of Reforms," *Canadian Journal of Economics* 30: 1234-47

[49] World Bank, 2005. *Economic Growth in the 1990s: Learning from a Decade of Reform*. Washington, DC.

[50] Xu, Chenggang. 2011. "The Institutional Foundations of China's Reforms and Development." *Journal of Economic Literature* 49: 1076-1151

[51] Zhang, Xiaobo and Dinghua Hu. 2011. "Overcoming Successive Bottlenecks: The Evolution of a Potato Cluster in China", *IFPRI* working paper.

Appendix 1. Proof of Lemma 1: By contradiction, suppose there exists an optimal adjustment time $t_1^* \in [0, \widehat{t})$ and a real adjustment is made so that $\delta_1 \neq \delta_0$. Substituting (8) into (9), we can easily prove that for $\forall t_1 \in (0, \widehat{t}]$,

$$\frac{\partial}{\partial t_1}\left\{\int_0^{t_1} h_0 e^{\mu t} e^{-rt} dt + e^{-rt_1}\left[V_0(\delta_1, h_0 e^{\mu t_1}) - C(\delta_0, \delta_1)\right]\right\} > 0, \forall \delta_1 \neq \delta_0.$$

Moreover, any adjustment affordable at t_1^* must be affordable at $\widehat{t}$. Contradiction. If it's optimal not to make any adjustment, $F_1(\delta_0, h_0) = V_0(\delta_0, h_0)$, where $t_1^* = \infty$. Q.E.D

Appendix 2. Proof of Lemma 2: The previous lemma shows $t_1^* \geq \widehat{t}$. When some nontrivial adjustment is made ($\delta_1^* < \delta_0 and T_1^* < \infty$), we must have $V_0(\delta_0, \frac{\eta}{\delta_0}) < V_0(\delta_1^*, \frac{\eta}{\delta_0}) - C(\delta_0, \delta_1^*)$. By Lemma 1 and the functional form of V_0 in (8), we must have

$$\frac{\partial\left[\text{RHS of } F_1(\delta_0, h_0)\right]}{\partial t_1} = re^{-rT_1}[V_0(\delta_0, \frac{\eta}{\delta_0}) + C(\delta_0, \delta_1) - V_0(\delta_1, \frac{\eta}{\delta_0})] \leq 0.$$

Since $t_1^* < \infty$, it must be that $\frac{\partial[\text{RHS of } F_1(\delta_0,h_0)]}{\partial t_1} < 0$ hence $t_1^* = \widehat{t}$.Q.E.D

Appendix 3. Proof of Proposition 1: We have the following two first order conditions with respect to t_1 and δ_1 :

$$\frac{\partial\left[\text{RHS of } F_1(\delta_0, h_0)\right]}{\partial t_1} = re^{-rt_1}[V_0(\delta_0, \frac{\eta}{\delta_0}) + C(\delta_0, \delta_1) - V_0(\delta_1, \frac{\eta}{\delta_0})] \leq 0$$
$$\frac{\partial\left[\text{RHS of } F_1(\delta_0, h_0)\right]}{\partial \delta_1} = 0 \Rightarrow \delta_1^* = \theta(\delta_0)\delta_0 \in [\eta, \delta_0) \text{ guaranteed by A1.}$$

The left inequality of (12) guarantees that $\delta_1^* < \delta_0$, while the right weak inequality of (12) makes sure that $\delta_1^* \geq \eta$. The condition $B < \widetilde{B}(\delta_0)$ ensures

$$V_0(\delta_0, \frac{\eta}{\delta_0}) + C(\delta_0, \delta_1^*) - V_0(\delta_1^*, \frac{\eta}{\delta_0}) < 0,$$

therefore $t_1^* = \widehat{t}$. The second order condition is also satisfied. Under A0 , (12), and $B < \widetilde{B}(\delta_0)$, we have

$$\begin{aligned} V_1(\delta_0, h_0) = & \frac{\mu\eta}{r(\mu - r)}\left(\frac{\eta}{\delta_0 h_0}\right)^{\frac{-r}{\mu}} \theta(\delta_0)^{\frac{r}{\mu}-1}\delta_0^{-1} \\ & - \frac{h_0}{\mu - r} - \left(\frac{\eta}{\delta_0 h_0}\right)^{\frac{-r}{\mu}} (A\theta(\delta_0)^{-\phi} + B), \end{aligned} \tag{35}$$

where $\theta(\delta_0)$ is given by (13). If $B > \widetilde{B}(\delta_0)$ or if the left inequality in (12) is violated, then $V_1(\delta_0, h_0) = V_0(\delta_0, h_0)$. If the right weak inequality in (12) is violated, then there are two possibilities. One is to waive the adjustment option because the adjustment cost dominates even the largest possible gain from an institutional adjustment, in which case we have

$$V_1(\delta_0, h_0) = V_0(\delta_0, h_0) \text{ if } V_0(\delta_0, \frac{\eta}{\delta_0}) + C(\delta_0, \eta) - V_0(\eta, \frac{\eta}{\delta_0}) \geq 0.$$

The second possibility is to exercise the adjustment option by fully exhausting the learning potential:

$$\delta_1^* = \eta \text{ and } V_0(\delta_0, \frac{\eta}{\delta_0}) + C(\delta_0, \eta) - V_0(\eta, \frac{\eta}{\delta_0}) < 0.$$

The second possibility requires $\widehat{B}(\delta_0) > B$, where

$$\widehat{B}(y) \equiv \frac{\mu}{r(\mu - r)} \left(\frac{\eta}{y}\right)^{\frac{r}{\mu}} - A\left(\frac{y}{\eta}\right)^{\phi} - \frac{\eta\mu}{r(\mu - r)y}.$$

In that case,

$$V_1(\delta_0, h_0) = \frac{\mu}{r(\mu - r)} (h_0)^{\frac{r}{\mu}} - \frac{h_0}{\mu - r} - \left(\frac{\eta}{\delta_0 h_0}\right)^{\frac{-r}{\mu}} (A(\frac{\eta}{\delta_0})^{-\phi} + B).$$

In summary, the value function with one adjustment opportunity is given by

$$V_1(\delta_0, h_0) = \begin{cases} \begin{array}{c} \frac{\mu\eta}{r(\mu-r)} \left(\frac{\eta}{\delta_0 h_0}\right)^{\frac{-r}{\mu}} \theta(\delta_0)^{\frac{r}{\mu}-1}\delta_0^{-1} \\ -\frac{h_0}{\mu-r} - \left(\frac{\eta}{\delta_0 h_0}\right)^{\frac{-r}{\mu}} (A\theta(\delta_0)^{-\phi} + B), \end{array} & \text{when} & \begin{array}{c} \widetilde{B}(\delta_0) > B \text{ and} \\ (12) \text{ is satisfied.} \end{array} \\ \begin{array}{c} \frac{\mu}{r(\mu-r)} (h_0)^{\frac{r}{\mu}} - \frac{h_0}{\mu-r} \\ -\left(\frac{\eta}{\delta_0 h_0}\right)^{\frac{-r}{\mu}} (A(\frac{\eta}{\delta_0})^{-\phi} + B), \end{array} & \text{when} & \begin{array}{c} \widehat{B}(\delta_0) > B \text{ and} \\ A\phi r \leq (A\phi r)^{\frac{1}{\phi + \frac{r}{\mu}}} < \frac{\eta}{\delta_0} \end{array} \\ V_0(\delta_0, h_0), & & \text{otherwise} \end{cases}.$$

Q.E.D.

Appendix 4. Proof of Lemma 3: The one-time control will be exercised non-trivially if and only if

$$V_0(\delta_0, \frac{\eta}{\delta_0}) < V_0(\delta, \frac{\eta}{\delta_0}) - C(\delta_0, \delta),$$

so $\delta < \delta_0$ only if

$$V_0(\delta_0, \frac{\eta}{\delta_0}) < V_0(\delta, \frac{\eta}{\delta_0}) - (A + B),$$

which implies (recall $\mu > r$)

$$\delta < \left[{\delta_0}^{\frac{r}{\mu}-1} + \frac{\delta_0^{\frac{r}{\mu}} (A + B)\, r(\mu - r)}{\mu\eta}\right]^{\frac{1}{\frac{r}{\mu}-1}}.$$

Since $\delta \geq \eta$, we require

$$\eta < \left[{\delta_0}^{\frac{r}{\mu}-1} + \frac{\delta_0^{\frac{r}{\mu}} (A + B)\, r(\mu - r)}{\mu\eta}\right]^{\frac{1}{\frac{r}{\mu}-1}},$$

or equivalently,

$$\left(\frac{\eta}{\delta_0}\right)^{\frac{r}{\mu}} > \frac{\eta}{\delta_0} + \frac{(A + B)\, r(\mu - r)}{\mu},$$

which is possible if and only if (16) is satisfied and $\frac{\eta}{\delta_0} \in (\underline{\beta}, \overline{\beta})$. (16) ensures $0 < \underline{\beta} < \left[\frac{\mu}{r}\right]^{\frac{1}{\frac{r}{\mu}-1}} < \overline{\beta} < 1$.Q.E.D.

Appendix 5. Proof of Proposition 2:

$$V_N(\delta_0, h_0) = \max\{G_N(\delta_0, h_0), F_N(\delta_0, h_0)\}, \tag{36}$$

where

$$\begin{aligned}
G_N(\delta_0, h_0) &\equiv \max_{t_1 \le \widehat{t}, \delta_1} \int_0^{t_1} h_0 e^{\mu t} e^{-rt} dt + e^{-rt_1} \left[V_{N-1}(\delta_1, h_0 e^{\mu t_1}) - C(\delta_0, \delta_1)\right]; \\
F_N(\delta_0, h_0) &\equiv \max_{\widehat{t} \le t_1, \delta_1} e^{-rt_1} \left[V_{N-1}(\delta_1, \frac{\eta}{\delta_0}) - C(\delta_0, \delta_1) - V_0(\delta_0, \frac{\eta}{\delta_0})\right] \\
&\quad + \frac{h_0 \mu}{\mu - r} \left(\frac{\eta}{\delta_0 h_0}\right)^{1-\frac{r}{\mu}} - \frac{h_0}{\mu - r}.
\end{aligned}$$

First observe

$$\begin{aligned}
G_N(\delta_0, h_0) &\equiv \max_{t_1 \le \widehat{t}, \delta_1} \int_0^{t_1} h_0 e^{\mu t} e^{-rt} dt + e^{-rt_1} V_{N-1}(\delta_1, h_0 e^{\mu t_1}) - e^{-rt_1} C(\delta_0, \delta_1) \\
&= \max_{t_1 \le \widehat{t}, \delta_1} \left[V_{N-1}(\delta_1, h_0) - e^{-rt_1} C(\delta_0, \delta_1)\right] \\
&\le \max_{\delta_1} \left[V_{N-1}(\delta_1, h_0) - e^{-r\widehat{t}} C(\delta_0, \delta_1)\right] \\
&\le F_N(\delta_0, h_0),
\end{aligned}$$

hence $V_N(\delta_0, h_0) = F_N(\delta_0, h_0)$ for any $N \ge 1$. That is, no adjustment will be made before the learning barrier becomes binding. Second, no adjustment will be made strictly after the learning barrier becomes binding. This is because $F_N(\delta, x) > V_{N-1}(\delta, x)$ if and only if there exists a $\widetilde{\delta} \in [\eta, \delta]$ such that

$$[V_{N-1}(\widetilde{\delta}, x) - C(\delta, \widetilde{\delta})] - V_{N-1}(\delta, x) > 0. \tag{37}$$

That is, an adjustment will be made if and only if the net benefit from the adjustment exceeds the value without adjustment. Suppose at time t the learning barrier becomes binding (that is, $x = \frac{\eta}{\delta}$), then if $F_N(\delta, x) > V_{N-1}(\delta, x)$, it's optimal to make the adjustment without any delay because of the time discounting. Note that the left hand side of (37) is the instantaneous value of net benefit from adjustment, which determines the optimal adjustment $\widetilde{\delta}$ by the first order condition. So the instantaneous value of the net benefit from adjustment is exactly the same for any time weakly after t. Moreover, the adjustment can always be fully financed because (37) implies the adjustment is profitable. The implied GDP dynamics is obvious.Q.E.D.

Appendix 6. Proof of Proposition 3: According to the previous proposition, there will be no adjustment after $\widehat{t} = -\frac{\ln h_0}{\mu}$, the time point when the developing country exactly achieves the same human capital level as the developed country if all the potential benefit of externality can be fully exploited. The total present discounted value of the gross benefit from the whole scheme of institutional adjustment can be no larger than

$$V_0(\eta, h_0) - V_0(\delta_0, h_0) = \frac{\mu}{r(\mu - r)} \left(\frac{1}{h_0}\right)^{\frac{-r}{\mu}} - \frac{\eta}{\delta_0} \frac{\mu}{r(\mu - r)} \left(\frac{\eta}{\delta_0 h_0}\right)^{\frac{-r}{\mu}}.$$

The present discounted cost of each downward adjustment can be no smaller than

$$e^{-r\widehat{t}}(A + B).$$

The minimum optimal number of adjustments is therefore no larger than $\frac{V_0(\eta,h_0)-V_0(\delta_0,h_0)}{e^{-r\widehat{t}}(A+B)}$. This is also true even when A or B equals zero.Q.E.D.

Appendix 7. Characterization of the problem when $N = 2$.

Similar to the previous case, when two adjustment opportunities are available, the value function becomes

$$V_2(\delta_0, h_0) = \max\{G_2(\delta_0, h_0), F_2(\delta_0, h_0)\},$$

where $G_2(\delta_0, h_0)$ is the value function when the first adjustment occurs before $\widehat{t}$:

$$G_2(\delta_0, h_0) \equiv \max_{t_1 \le \widehat{t}, \delta_1} \int_0^{t_1} h_0 e^{\mu t} e^{-rt} dt + e^{-rt_1} \left[V_1(\delta_1, h_0 e^{\mu t_1}) - C(\delta_0, \delta_1)\right],$$

and $F_2(\delta_0, h_0)$ is the value function when the first adjustment occurs after $\widehat{t}$:

$$F_2(\delta_0, h_0) \equiv \max_{\widehat{t} \le t_1, \delta_1} \left[\begin{array}{c} \int_0^{\widehat{t}} h_0 e^{\mu_0 t} e^{-rt} dt + \int_{\widehat{t}}^{t_1} \frac{\eta}{\delta_0} e^{-rt} dt \\ +e^{-rt_1} \left[V_1(\delta_1, \frac{\eta}{\delta_0}) - C(\delta_0, \delta_1)\right] \end{array} \right].$$

Note that

$$\begin{aligned} G_2(\delta_0, h_0) &= \max_{t_1 \le \widehat{t}, \delta_1} \left[V_1(\delta_1, h_0) - e^{-rt_1} C(\delta_0, \delta_1)\right] \\ &\le \max_{\delta_1} \left[V_1(\delta_1, h_0) - e^{-r\widehat{t}} C(\delta_0, \delta_1)\right] \le F_2(\delta_0, h_0), \end{aligned}$$

therefore $V_2(\delta_0, h_0) = F_2(\delta_0, h_0)$. Suppose two nontrivial adjustments are made. There are two possibilities. First, when (12) is satisfied, the first order condition with respect to δ_1 yields

$$B\frac{r}{\mu}\delta_1^{\frac{r}{\mu}+\phi} = A\delta_0^{\phi+\frac{r}{\mu}}\phi - k\delta_1^{\frac{r}{\mu}+\phi-\frac{\phi}{\frac{r}{\mu}+\phi-1}}, \tag{38}$$

where $k \equiv \frac{\eta(\phi\mu+r)}{r\phi\mu}\left(\frac{Ar\phi}{\eta}\right)^{\frac{\frac{r}{\mu}-1}{\frac{r}{\mu}+\phi-1}}$. No closed-form solution can be obtained, but it can be shown that the solution exists and is unique if $\frac{r}{\mu} + \phi \ge 2$, which is assumed true. In Figure A, the upward-sloping curve plots the term on the left hand side of (38) while the downward-sloping curve corresponds to the right hand side.

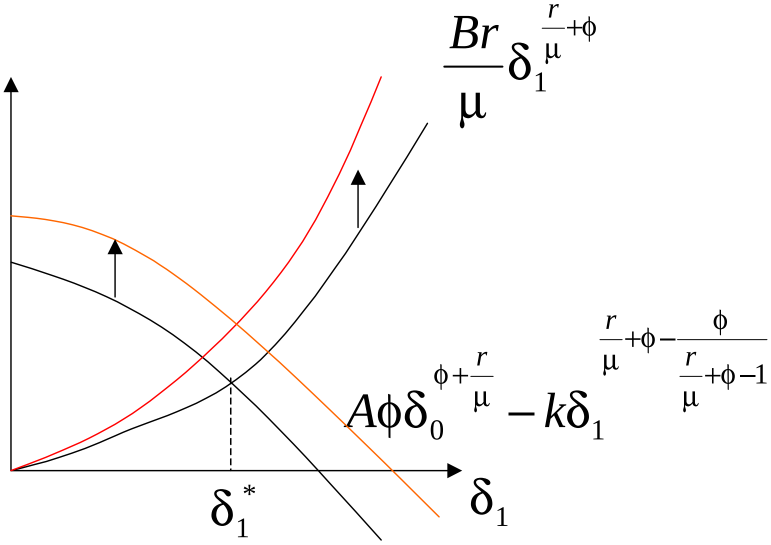

Figure A. Optimal Adjustment Size when $N = 2$

Let δ_1^* denote the unique solution to equation (38). $\delta_1^* > \eta\left(Ar\phi\right)^{-\frac{1}{\frac{r}{\mu}+\phi}}$ must hold, so

$$\eta < \delta_0 \left[\frac{A^2 r\phi^2}{B\frac{r}{\mu} + \frac{(\phi\mu+r)}{r\phi\mu}\left(Ar\phi\right)^{\frac{\left(\frac{r}{\mu}-1\right)\left(\frac{r}{\mu}+\phi\right)+\phi}{\left(\frac{r}{\mu}+\phi-1\right)\left(\frac{r}{\mu}+\phi\right)}}}\right]^{\frac{1}{\frac{r}{\mu}+\phi}}. \tag{39}$$

In addition, $\delta_1^* < \delta_0$ must also hold, or equivalently,

$$A\phi - B\frac{r}{\mu} \leq 0,\ \text{or}\ \eta < \delta_0 \left[\frac{r\phi\mu\left(A\phi - B\frac{r}{\mu}\right)}{\left(\phi\mu+r\right)\left(Ar\phi\right)^{\frac{\frac{r}{\mu}-1}{\frac{r}{\mu}+\phi-1}}}\right]^{\frac{\frac{r}{\mu}+\phi-1}{\phi}} \quad \text{when } A\phi - B\frac{r}{\mu} > 0. \tag{40}$$

Moreover, $t_1^* = \widehat{t}$ if $V_0(\delta_0, \frac{\eta}{\delta_0}) + C(\delta_0, \delta_1^*) - V_1(\delta_1^*, \frac{\eta}{\delta_0}) \leq 0$, which is equivalent to $\Delta(\delta_0, \delta_1^*) \geq 0$ when both (39) and (40) are satisfied, where

$$\begin{aligned}\Delta(\delta_0, \delta_1^*) \equiv {} & (\delta_1^*)^{\frac{r}{\mu}-\frac{\phi}{\frac{r}{\mu}+\phi-1}}\left(\frac{Ar\phi}{\eta}\right)^{\frac{\frac{r}{\mu}-1}{\frac{r}{\mu}+\phi-1}}\frac{\eta}{r\phi}\left(\frac{\phi\mu}{\mu-r}-1\right)\delta_0^{-\frac{r}{\mu}} \\ & -\frac{\eta\mu}{r(\mu-r)\delta_0} - B - B\left(\frac{\delta_1^*}{\delta_0}\right)^{\frac{r}{\mu}} - A\delta_0^{\phi}\delta_1^{*-\phi}.\end{aligned}$$

We still need to check whether $\widetilde{B}(\delta_1^*) \geq B$. When $\Delta(\delta_0, \delta_1^*) \geq 0$ and $\widetilde{B}(\delta_1^*) \geq B$ are both satisfied, we have $\delta_2^* = \theta(\delta_1^*)\delta_1^*$ and $t_2^* = \frac{1}{\mu}\ln\frac{\eta}{\delta_1^* h_0}$. The developing economy grows at speed $(\mu + g_H)$ up to the time point $\frac{1}{\mu}\ln\frac{\eta}{\delta_2^* h_0}$, after which the convergence stops and there will be a permanent GDP gap between the two economies ($\frac{\eta}{\delta_2^*} < 1$).

Comparative statics analysis shows the following: An increase in B will move δ_1^* leftward (see Figure A), this is because the costing-saving motive will make adjustment less frequent but the size for each adjustment larger. In contrast, a higher A results in a higher δ_1^* because the variable adjustment cost parameter A affects the marginal adjustment cost. The higher the initial barrier, the more modest the first target of barrier adjustment.

In addition, we have $\frac{\partial \delta_1^*}{\partial \eta} < 0$, implying that the first adjustment is larger when the relative scale of the economy becomes bigger and $\frac{\partial \delta_1^*}{\partial \delta_0} > 0$, meaning the institutional barrier exhibits certain persistence as an initial inferior institution leads to an relatively inferior institution after the first adjustment.

When $\widetilde{B}(\delta_1^*) < B$ or $\Delta(\delta_0, \delta_1^*) < 0$ or any other conditions are not satisfied, the following problem needs to be solved:

$$F_2(\delta_0, h_0) \equiv \max_{\delta_1} \int_0^{\widehat{t}} h_0 e^{\mu_0 t} e^{-rt} dt + e^{-r\widehat{t}} \left[V_1(\delta_1, \frac{\eta}{\delta_0}) - C(\delta_0, \delta_1) \right],$$

where

$$V_1(\delta_1, \frac{\eta}{\delta_0}) = \frac{\mu}{r(\mu - r)} \left(\frac{\eta}{\delta_0} \right)^{\frac{r}{\mu}} - \frac{\eta}{\delta_0 (\mu - r)} - \left(\frac{\delta_0}{\delta_1} \right)^{\frac{-r}{\mu}} (A(\frac{\eta}{\delta_1})^{-\phi} + B).$$

The first order condition is

$$-(\frac{r}{\mu} + \phi) A \eta^{-\phi} \delta_1^{\phi} + \phi A \delta_0^{\phi + \frac{r}{\mu}} \delta_1^{-\phi - \frac{r}{\mu}} = B \frac{r}{\mu}, \tag{41}$$

which implies the existence and uniqueness of the root δ_1^*. So $\frac{\partial \delta_1^*}{\partial \eta} > 0$; $\frac{\partial \delta_1^*}{\partial \delta_0} > 0$; $\frac{\partial \delta_1^*}{\partial A} > 0$; $\frac{\partial \delta_1^*}{\partial B} < 0$. The only difference from the previous case is that $\frac{\partial \delta_1^*}{\partial \eta}$ has a different sign. The following two conditions need verifying:

$$V_1(\delta_1^*, \frac{\eta}{\delta_0}) \geq C(\delta_0, \delta_1^*) + V_0(\delta_0, \frac{\eta}{\delta_0}), \tag{42}$$

and $\widehat{B}(\delta_1^*) \geq B$. (41) implies that $\delta_1^* > \eta$ is equivalent to $\eta < \left[\frac{A\phi}{(A+B)\frac{r}{\mu} + A\phi} \right]^{\frac{1}{\frac{r}{\mu} + \phi}} \delta_0$. And to ensure $\delta_2^* = \eta$, we must require $\eta \geq \left[\frac{A^2 \phi^2 r}{B\frac{r}{\mu} + A(\phi + \frac{r}{\mu})(A\phi r)^{-\frac{\phi}{\frac{r}{\mu}+\phi}}} \right]^{\frac{1}{\frac{r}{\mu} + \phi}} \delta_0$. It also ensures $\delta_2^* < \delta_1^*$. To summarize, we have the following result:

Suppose $N = 2$ and both Assumptions A0 and (12) are satisfied. [1] $\delta_2^* = \eta$ if and only if $\left[\frac{A^2 \phi^2 r}{B\frac{r}{\mu} + A(\phi + \frac{r}{\mu})(A\phi r)^{-\frac{\phi}{\frac{r}{\mu}+\phi}}} \right]^{\frac{1}{\frac{r}{\mu} + \phi}} \leq \frac{\eta}{\delta_0} < \left[\frac{A\phi}{(A+B)\frac{r}{\mu} + A\phi} \right]^{\frac{1}{\frac{r}{\mu} + \phi}}$, (42) is satisfied and $\widehat{B}(\delta_1^*) \geq B$, where δ_1^* is uniquely determined in (41); [2] $\delta_2^* > \eta$ if and only if $\Delta(\delta_0, \delta_1^*) \geq 0$, $\widetilde{B}(\delta_1^*) \geq B$,

(39) and (40) are all satisfied, where δ_1^* is uniquely determined by (38) and $\delta_2^* = \theta(\delta_1^*)\delta_1^*$;
[3] Otherwise, $V_2(\delta_0, h_0) = V_1(\delta_0, h_0)$ given by (14).

Appendices 8. Proof of Proposition 4: The first order condition with respect to δ_i is

$$\left(\frac{\delta_i}{\delta_{i+1}}\right)^{\phi}\left[\frac{r}{\mu}+\phi\right]+\frac{B}{A}\frac{r}{\mu}=\left[\frac{\delta_{i-1}}{\delta_i}\right]^{\frac{r}{\mu}+\phi}\phi \text{ for } \forall i < N^*, \tag{43}$$

which can be rewritten as

$$\left(\frac{y_{i+1}}{y_i}\right)^{\phi}=\frac{y_i^{\frac{r}{\mu}}\phi-\frac{B}{A}\frac{r}{\mu y_i^{\phi}}}{\frac{r}{\mu}+\phi} \text{ for } \forall i < N^*,$$

which requires $y_i > \left[\frac{rB}{A\mu\phi}\right]^{\frac{1}{\phi+\frac{r}{\mu}}}$. It also implies

$$\frac{y_{i+1}}{y_i} \gtreqless 1, \forall i \in \mathcal{N}, \text{ iff } y_j \gtreqless \varpi, \text{ for some } j \in \mathcal{N},$$

where ϖ is uniquely determined by

$$\omega^{\frac{r}{\mu}}\phi-\frac{B}{A}\frac{r}{\mu\varpi^{\phi}}=\frac{r}{\mu}+\phi.$$

The first order condition with respect to δ_{N^*} is

$$\begin{aligned}\left(\frac{Ar\phi}{\eta}\delta_{N-1}^{\frac{r}{\mu}+\phi}\right)^{\frac{1}{\frac{r}{\mu}+\phi-1}} &= \delta_N \text{ if } \delta_N > \eta, \\ \left(\frac{Ar\phi}{\eta}\delta_{N-1}^{\frac{r}{\mu}+\phi}\right)^{\frac{1}{\frac{r}{\mu}+\phi-1}} &\le \delta_N \text{ if } \delta_N = \eta.\end{aligned} \tag{44}$$

To solve the problem completely, we define $\delta_{N-2}^* \equiv \Gamma(\delta_{N-1}^*, \delta_N^*)$ from (43) when $\delta_N^* > \eta$.

Obviously, $\Gamma_1 > 0$ and $\Gamma_2 < 0$. Recursively, we have

$$\begin{aligned}&\delta_{N-3}^* = \Gamma(\delta_{N-2}^*, \delta_{N-1}^*) = \Gamma(\Gamma(\delta_{N-1}^*, \delta_N^*), \delta_{N-1}^*);\\ &\delta_{N-4}^* = \Gamma(\delta_{N-3}^*, \delta_{N-2}^*) = \Gamma(\Gamma(\Gamma(\delta_{N-1}^*, \delta_N^*), \delta_{N-1}^*), \Gamma(\delta_{N-1}^*, \delta_N^*));\\ &\quad ...\end{aligned}$$

We ultimately have δ_0 as a function of δ_{N-1}^* and δ_N^*. Together with (44), both δ_{N-1}^* and δ_N^* , hence everything else, can be pinned down. When $\delta_N = \eta$, we have $\delta_{N-2}^* \equiv \Gamma(\delta_{N-1}^*, \eta)$. Using the same recursive substitution, we can express δ_0 as a function of δ_{N-1}^*, from which δ_{N-1}^* hence δ_i^* can be obtained for $\forall i = 1, 2, ...N$.

[1] When $A = 0$, (19) becomes

$$\max_{N,\{\delta_i\}_{i=1}^N}\frac{\mu\eta}{r(\mu-r)}\left(\delta_N\right)^{\frac{r}{\mu}-1}-B\sum_{i=1}^{N}\delta_{i-1}^{\frac{r}{\mu}}.$$

Obviously, $N^* = 1$ if any adjustment will be made. $\delta_1^* = \eta$ is the solution to the following problem

$$\max_{\delta_1 \geq \eta} \frac{\mu\eta}{r(\mu - r)} (\delta_1)^{\frac{r}{\mu}-1} - B\delta_0^{\frac{r}{\mu}}.$$

[2] When $A \neq 0$, Lemma 3 implies that $\delta_N = \theta(\delta_{N-1})\delta_{N-1}$ when $\delta_N > \eta$. Thus the reform sizes are strictly increasing if and only if $y_N > \varpi$, or equivalently, $\delta_N < \frac{\eta}{A\phi r}\varpi^{-\left(\phi+\frac{r}{\mu}\right)}$. It means that the long run GDP $h = \frac{\eta}{\delta_N} > A\phi r \varpi^{\phi+\frac{r}{\mu}}$.

When $\delta_{N^*}^* = \eta$, and $N^* \geq 2$, then by applying (41), we have

$$-(\frac{r}{\mu} + \phi)A\left[\frac{\delta_{N^*-1}}{\eta}\right]^{\phi} + \phi A\left[\frac{\delta_{N^*-2}}{\delta_{N^*-1}}\right]^{\phi+\frac{r}{\mu}} = B\frac{r}{\mu},$$

then equal adjustment size is possible only if

$$\frac{\delta_{N^*-1}}{\eta} = \frac{\delta_{N^*-2}}{\delta_{N^*-1}} = \left[\frac{(\frac{r}{\mu} + \phi)}{\phi}\right]^{\frac{\mu}{r}},$$

which is still consistent with the optimal size obtained when $\delta_{N^*}^* > \eta$. Q.E.D.

Appendix 9. Proof of Proposition 5: Under assumption (12) and $B < \widetilde{B}\ (\delta_0)$, the optimal adjustment size given by (13) can be fully financed by the domestic saving of the developing economy if and only if $Q(h_0) \geqslant 0$, where

$$Q(z) \equiv \frac{z}{(\mu - r)}\left[\frac{\eta}{\delta_0 z} - \left(\frac{\eta}{\delta_0 z}\right)^{\frac{r}{\mu}}\right] - (A\left[\frac{A\phi r\delta_0}{\eta}\right]^{\frac{-\phi}{\phi+\frac{r}{\mu}-1}} + B).$$

Obviously, $Q(z) < 0$ for any z when (29) is violated. In that case, the first-best reform is never feasible. Observe that $Q'(z) < 0$ whenever $z > 0$. There is a unique root, h^*,of the equation $Q(h^*) = 0$. It can be verified that $h^* < \frac{\eta}{\delta_0}$ because of (29). Thus $Q(h_0) > 0$ iff $h_0 < h^*$.