2010

Essays on optimal allocation of resources by governments

Monisankar Bishnu
Iowa State University

Essays on optimal allocation of resources by governments

by

Monisankar Bishnu

A dissertation submitted to the graduate faculty
in partial fulfillment of the requirement for the degree of
DOCTOR OF PHILOSOPHY

Major: Economics

Program of Study Committee:
Joydeep Bhattacharya, Co-Major Professor
Krishna B. Athreya, Co-Major Professor
Helle Bunzel
Arne Hallam
Rajesh Singh

Iowa State University
Ames, Iowa
2010

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

Words are not enough to express my sense of gratitude towards Professor Joydeep Bhattacharya. He has always been a great inspiration to me. His clarity of economic thought and unique approach to a problem was a great stimulation. Whatever little I have learned from him will definitely be an asset for the rest of my life. I am deeply indebted to him. I have also been extremely fortunate to get the opportunity to work under the guidance of Professor Krishna B. Athreya. His support and encouragement were a great motivation behind the completion of my doctoral thesis. I would like to express my sincere gratitude to him. I believe that whatever I have been able to learn from him will help me immensely in my future endeavors. I would like to thank Professor Helle Bunzel for her extremely useful suggestions on my research. I am indebted to Professor Arne Hallam for his generous suggestions on my research; without his help, my research would not have been completed. I am grateful to Professor Rajesh Singh with whom I have had many useful discussions on different issues. I admire his encouragement, enthusiasm and humor.

I am indebted to my parents, Tara Sankar and Gita Bishnu, for their continuous support and encouragement. The inspiration, patience and devotion from my wife Sonali were invaluable. I must also thank my elder brother Dip Sankar, without whose encouragement I would not have ventured into economics. My little daughter Beas entertained me whenever I needed a break from studying. I also owe a lot to my extended family members and friends.

CHAPTER 1. GENERAL INTRODUCTION

Allocation of resources is a fundamental issue in economics. My thesis is on how the government can optimally allocate the available resources so that the economy can exploit its benefits to the fullest extent. In Chapter 2, the issue of how labor resources can be optimally allocated across different governmental agencies is discussed. The main focus is to study how the efficiency of government activities depends on the allocation of labor resources by the government. This analysis is very time appropriate and crucial from the efficiency point of view. More specifically, the idea of this chapter is as follows. In many countries, the process of obtaining government approval for different projects involves interaction with multiple government agencies at various levels. This often makes the approval process inefficient by unnecessary lengthening it. In this paper we study the effect of a re-organization of the approval process towards making it a single window clearance system, on the efficiency of the entire process.

In Chapter 3, the rationale behind government intervention in education in the presence of consumption externalities is studied. Government intervention in education, typically in the form of education subsidies, is ubiquitous. The standard rationale for such intervention is a human capital externality: people are smarter if they are around smart people. The intergenerational counterpart of this observation is that a smarter generation produces a smarter future generation. This paper argues for government intervention in education even when no human capital externalities are present. To that end, a neoclassical overlapping generations model of human and physical capital accumulation is studied. Children borrow from perfect capital markets to fund education expenses. When middle-aged, they earn income from human capital, and save in the form of physical capital. Agents are assumed to care not just about the level of their consumption but how that compares to those (from among their peers and the other consuming generation) living around them. Such an intra and intergenerational consumption externality is responsible for the possibility that agents may over or under-accumulate human and physical capital relative to a planner, thereby justifying public intervention.

Finally, Chapter 4 investigates how an economy with production shock behaves in the long run under the presence of government subsidy in education. A growth model is considered where human capital in each period is generated using the human capital available from the previous generation as well as an education subsidy. The education subsidy is made available through a generational transfer mechanism where tax is collected from the working class and passed on to the new generation. The final good in the economy is produced using both physical and human capital. However, the production process is subject to a periodwise shock. In this stochastic framework we define a balanced growth path where the distribution of the ratio of two capital stock converges. The conditions under which this invariant measure exists and is unique are completely specified. The tax rate for which this benchmark economy follows the same growth path as a complete market has been studied. It has also been shown that the invariant distribution for an economy with a tax rate lower than this First Order Stochastically Dominates the invariant distribution of a complete market.

CHAPTER 2. ON THE EFFICIENCY OF SINGLE WINDOW

Modified from a paper published in Economic Theory

Krishna B. Athreya and Monisankar Bishnu

1 Introduction

In the last decade or so, growing economies like India are emerging as important players in the global arena. Globalization has led to an increased interaction with the world. There is a growing willingness among foreign investors to invest in India, just as the number of Indian citizens eager to conduct business with foreign countries is increasing. All this is well because it leads to an overall increase in economic activities which in turn triggers higher growth of a country. But what has been observed in countries like India is that the sluggishness and inefficiency of the government bureaucracy is a major hurdle in the path of development. Indeed, as observed by Athreya and Majumdar (2005), this feature is shared by many developing countries. Our particular area of focus is the relative inefficiency of the system of *multiple approvals* which is in place in many developing countries. To be more elaborate, increase in economic activity within a country leads to more people applying for passports, investors applying for approvals for new projects, potential exporters applying for a license etc. They have to approach the government bureaucracy, which is the sole custodian of granting approvals.

An investment proposal, for example might need to be assessed from various angles. While it is necessary to ascertain whether the project is financially viable, it might also be important to confirm whether the project poses any environmental hazard. It is natural to expect that the people with the expertise to judge the financial aspects of an investment project may not be the best ones to assess its environmental implications. Hence, an investor has to approach different bureaucrats in different establishments for approval before being able to go ahead with the project. This system is technically referred to as *multiple windows*. However this decision making process of granting approvals, which comprises of "a series of substantially meaningless scrutinies" (Bhagwati 1973), is generally very slow. This is largely due to a lack of coordination among the different government agencies involved. For example, in order to be able to export, an Indian citizen needs to acquire 257 signatures and 118 copies of the same document. This involves dealing with different Government of India agencies separately. The magnitude of the delay resulting in the process is reflected by the fact that 118 copies of the same document need 22 hours to key punch! (See Roy 2003; Athreya and Majumdar 2005).

The World Bank reports 'Enterprise surveys' (2008) and 'Doing Business' (2008) reveal that most developing countries have not been able to create an environment where it is easy to conduct business. According to the 'Doing Business' report, which ranks countries according to the 'ease of doing business' there, out of 178 countries, the ranks of China, India and Brazil are 83, 120 and 122 respectively. As an example, getting various approvals for building a warehouse takes on an average 224 days in India, 336 days in China and 411 days in Brazil, whereas the average in US and Canada are 40 and 75 days respectively. According to 'Enterprise surveys' report, percentage of senior management time spent dealing with the requirements of government regulation is 6.66% for India, 7.19% for Brazil, 18.3% for China and 20.49% for Mexico whereas the OECD average is only 2.97%. The delay caused by having to deal with different agencies or '*windows*' often results in the country being deprived of good investment, as some foreign investors frustrated by the sluggishness and opaqueness of the process decide to take their project elsewhere. Let us define social welfare as the sum of the benefits from different projects. If the benefit from a project is the cost of waiting for approval in the queue in the steady state subtracted from the gain from approvals, then one of the decisive factors in the measurement of social welfare is the waiting cost which in turn depends on the waiting time at the steady state. Thus, a long waiting time in the queue has the undesirable effect of lowering social welfare. To counter the negative impacts of a *multiple window* approval system, policy makers emphasize the importance of a *single window*[1].

In this paper we show that single window clearance is more efficient than the prevailing systems of multiple approvals, where efficiency is defined in terms of both expected queue size and expected waiting time in a stochastic equilibrium. Not only that, for some parameter values, we find that as the rate of incoming applications increases, the gain in efficiency achieved under a single window clearance system over all other systems, in the steady state, actually increases[2]. This observation is all the more significant for emerging countries like India, where the number of people willing to invest (and hence the number of 'applications') is on the rise. Hence the case for adopting the single window system is even stronger.

It is not as if the government is unaware of the need to implement a single window clearance system. In spite of this, not much progress has been made in this direction.

[1] According to the Recommendation No. 33 (New York and Geneva 2005) of 'Recommendation and Guidelines on Establishing a Single Window', prepared by the United Nations Center for Trade Facilitation(UN/CEFACT), a) as the single window enables governments to process submitted information, documents and fees both faster and more accurately, traders should benefit from faster clearance and release times, enabling them to speed up the supply chain. In addition, the improved transparency and increased predictability can further reduce the potential for corrupt behaviour from both the public and private sector. b) Major benefit for the government inludes 'more effective and efficient deployment of resources' and 'increased integrity and transparency'.

[2] However, we suspect that this result holds for all possible values of the parameters. The result can be shown to hold easily when the service rates of the two governments are equal.

One of the main reasons that is cited for this lack of an affirmative action towards this end is corruption in the government circles, and an unholy nexus between the politicians and bureaucrats. The research on single window policy has gained momentum very recently and one of the first of its kind is Athreya and Majumdar (2005). More recently, Lambart-Mogiliansky, Majumdar and Radner (henceforth L-M-R 2007, 2008) and Yoo (2007) presented a game theoretic approach to analyze single window policy in the presence of corruption. However, there might be another possible explanation for this inaction. The age old structure of the government in countries like India, typically overemploys labor. It is sometimes the fear in the government circles and a general public perception that any re-organization might result in loss of government jobs, that prevents any populist government from taking any drastic decision. Unemployment is a very sensitive issue for any populist government in a developing country. Any decision that might be generally perceived to have a negative effect on employment may hamper the chances of that government coming back to power in the future. This breeds and nourishes inefficiency in the approval process. We claim that the re-organization required to increase the efficiency of the approval system does not have to result in any loss of jobs. We present a simple queue theoretic framework to present our claims. Thus in order to boost the growth progress, it is the duty of government to raise the level of awareness among its own officials as well as the general public. This requires a strong political will to change the way of functioning of the government bureaucracy, and the sooner it comes the better.

The rest of the paper is organized as follows. In section 2, we present our model in an $M/M/(.)$ framework and subsequently in the subsections 2.1 and 2.2, we show that under two different measures of inefficiency, a single window scores better than the other existing approval systems. In subsection 2.3, we investigate the relationship between the inefficiencies of the different approval systems and the arrival rate as well as the service rate at the steady state for different parameter values. In section 3, we compare different structures of a single window clearance system and justify why the extreme version of the single window setup has been chosen for our discussion in section 2. In subsection 3.1, we present a simple $M/M/(.)$ framework to produce our results while in subsection 3.2, a more general $M/G/(.)$ framework has been considered. Section 4 concludes.

2 The Framework

In this section we discuss three different systems of approval, and evaluate them in terms of their efficiency. As noted previously, under a multiple window approval system, the applicant has to deal with a number of separate agencies like the local governments, the federal government, the central bank etc. In this paper, we assume for the sake of simplicity that under a multiple window approval system, the applicant has to deal with two separate agencies, namely the local government and the federal government (henceforth referred to as LG and FG respectively). We also assume that all applications are homogeneous, i.e., one application is no way differentiable from the other. Furthermore, both FG and LG are assumed to have the capacity to receive infinite number of applications. The three systems of approval that we discuss are as follows:

Simultaneous approval system (SIS): Under this system, an application is approved if and only if it is passed independently by both FG and LG. This might happen in cases where the project involves issues which lie under the jurisdiction of both the local and the federal governments. Under SIS, arrivals of applications to both the governments are Poisson processes with rate λ. The service process of the FG and LG are also Poisson processes with rate μ_f and μ_l respectively, where subscript f stands for the FG and l

for the LG. The arrival process of applications to FG is assumed to be independent of the arrival process of applications to LG. The service process of applications of FG is also assumed to be independent of the service process of applications of LG, i.e., the service times are mutually independent random variables with distribution functions $exponential(\mu_f)$ and $exponential(\mu_l)$ respectively. Furthermore, the arrival process of applications is independent of the service process for both FG and LG.

Sequential approval system (SES): Under this system, a project has to be first approved by the FG and then by the LG in order to be executed. Thus, an application comes before the LG for approval only if it has already been approved by the FG. This sequential form of clearance is also a common form of governmental approval process in some countries. Under SES, the arrival process to the first window, here the FG, is a Poisson process with rate λ. The service rates of both FG and LG are exactly identical to that in SIS. Just as in SIS, the service process of applications of FG is assumed to be independent of the service process of applications of LG, i.e., the service times are mutually independent random variables with distribution functions $exponential(\mu_f)$ and $exponential(\mu_l)$ respectively. The arrival process of applications is also independent of the service process for FG.

Single Window approval system (SWS): Under this system, an applicant has to deal with only a single establishment in order to get approval for the project. This single window approval processing institution may be a separate agency[3] as well. Here, the

[3] According to the 'Recommendation and Guidelines on Establishing a Single Window' to enhance the efficienct exchange of information between trade and government, Recommendation No. 33 (New York and Geneva 2005), prepared by the United Nations Center for Trade Trade Facilitation(UN/CEFACT), 'The appropriate agency to lead the establishment and operation of a Single Window will vary from country to country depending on legal, political and organisational issues. The lead agency must be a very strong organisation with the necessary vision, authority (legal), political backing, financial and human resources and interfaces to other key organisations. However, the lead organisation does not necessarily have to be a governmental organisation; it can be a private entity such as a Chamber of Commerce or a semi-state organisation such as a Board of Trade. However, private organisations sometimes lack the legal authority to issue and accept information and documents and the

arrival process to the single window is Poisson with rate λ and the service process is also Poisson with rate $[\frac{1}{2}(\mu_f^{-1} + \mu_l^{-1})]^{-1}$. Under SWS we assume that the arrival process of applications to the single window is independent of the service process.

The justification behind choosing this particular rate for SWS is as follows. The arrival rate of applications is λ, because one can think of applications under this policy to be a 'one single complex license' (see for e.g. L-M-R 2007) that takes care of everything. However, each application requires exponentially distributed amount of time with mean μ_f^{-1} from the FG and μ_l^{-1} from LG. Therefore, a single window can process applications at a rate $[\frac{1}{2}(\mu_f^{-1} + \mu_l^{-1})]^{-1}$ by appropriately using its labor force. For example, if the service time of approval by both the FG and LG are exponentially distributed and the mean service time to process their parts of an application is say one month for both the governments, then we argue that to process a full application, the mean service time required by a combined approval under SWS is simply one month. This rate of service can be achieved by re-organizing the workloads on the existing employees of the FG and LG, without affecting the total employment level. Further, a SWS also eliminates the problem of lack of coordination that often exists between two different government agencies, to some extent. It also seems logical to assume that when there is a gap between the efficiencies, working with the more efficient labor may help improving the quality of the less efficient labor force. Thus we expect that under the SWS, a single agency may have some efficiency gain through 'effective and efficient deployment of resources' (see footnote 1) when all the processes take place under a single umbrella. Later we also carry out our analysis under the assumption that the expected service

power to enforce rules. Therefore, in such a scenario, it may be necessary for the private organisation to seek the explicit formal support of a governmental organisation that has such power at its disposal.

One example of a public-private partnership that led to the establishment of a Single Window was the Mauritius Network Services Ltd in Mauritius. This is a tripartite joint-venture company involving public and private sector representatives and a foreign technical partner.'

time of SWS is the maximum of the expected service time (hence the minimum of the service rates) of the two governments. That is, if the service time of the FG and LG are exponentially distributed and the mean service time required to process an application is say one month and two months respectively, then we assume that a SWS has a mean service time of two months, which is the maximum of the expected service times of the two governments. We show that the change in the assumption about the service time of a SWS does not lead to any change in our result. The structure of the model ensures that any possible re-organization of the bureaucracy due to a change in the approval system does not result in any unemployment.

For both SIS and SES, define $\rho_i = \frac{\lambda_i}{\mu_i}$, $i = f, l$ as the occupation rate. For the existence of the steady state, we assume that $\rho_i < 1$; otherwise the mean queue length explodes. For SWS, define the occupation rate as $[\frac{1}{2}(\mu_f^{-1} + \mu_l^{-1})]\lambda$ and for existence of the steady state, we assume that $[\frac{1}{2}(\mu_f^{-1} + \mu_l^{-1})]\lambda < 1$.

In the following subsections we present two different measures of inefficiency. We then evaluate SIS, SES and SWS according to these measures. In subsection 2.1, we present the first measure of inefficiency namely the *expected number of applications* in the system at the steady state (henceforth ENA). The second measure, which is the *expected waiting time* of an application in the steady state (henceforth EWT), is presented in subsection 2.2. The reason behind choosing these inefficiency measures is that if the length of the queue in the steady state is long, it essentially implies a potentially greater loss of productive activity. A longer mean waiting time in the steady state also has the same implication on social welfare.

2.1 Inefficiency under ENA

A stochastic equilibrium, i.e., an invariant or steady state distribution $\pi^* = (\pi_n^*)$ of the queue size exists for each of the queues defined in SIS. Each of the queues in SIS is $M/M/1$ in the language of Queueing Theory. From the above description, it is clear that the transition rates for each of the queues under SIS are $q_i(n, n{+}1) = \lambda$ and $q_i(n, n{-}1) = \mu_i$; $i = f, l$ where $q_i(n, n+1)$ and $q_i(n, n-1)$ are the birth and death rates when there are n applications in the system. Therefore the form of the steady state distribution for each of the queues in SIS is given explicitly by $\pi_n^* = \Pr[X = n] = (\frac{\lambda}{\mu})^n(1-\frac{\lambda}{\mu}) = \rho^n(1-\rho)$ (Durrett 1999; Klienrock 1975). Note that given the irreducibility, π^* is approximately the distribution of the queue size for large t, i.e., $\pi_n^* = \lim_{t\to\infty} \Pr[X(t) = n]$ where $X(t)$ is the number of applications in one queue at time t. Furthermore, in the steady state, the average number of applications in the system, $\lim_{t\to\infty} \frac{1}{t}\int_{y=0}^{t} X(y)dy = EX$ holds w.p. 1. So under SIS, the expected number of applications in the system for each of the two queues is given by

$$EX^i = \sum_{n=0}^{\infty} n(\rho_i)^n(1-\rho_i) = \frac{\rho_i}{(1-\rho_i)} \ , \quad i = f, l$$

We are now interested to find out the maximum length of the two queues under SIS at the steady state. This measure is chosen because even if the number of steady state applications still waiting for approval from one government is low, applicants will have to wait longer because of the other relatively less efficient government. We measure the inefficiency of SIS by the expected number of applications still awaiting approval in the steady state and denote it by $ENA(SIS)$. Given that the applications are homogeneous and the two processes are independent, we have

$$\begin{aligned}
\Pr[\max\{X^f, X^l\} &= n] = \Pr(X^f = n, X^l = n) + \Pr(X^f = n, X^l \leqq n-1) \\
+\Pr(X^l &= n, X^f \leqq n-1) \\
&= \rho_f^n \rho_l^n (1-\rho_f)(1-\rho_l) + \sum_{m=0}^{n-1} \rho_f^n (1-\rho_f) \rho_l^m (1-\rho_l) \\
&\quad + \sum_{m=0}^{n-1} \rho_l^n (1-\rho_l) \rho_f^m (1-\rho_f) \\
&= \rho_f^n \rho_l^n (1-\rho_f)(1-\rho_l) + \rho_f^n (1-\rho_f)(1-\rho_l^n) + \rho_l^n (1-\rho_l)(1-\rho_f^n)
\end{aligned}$$

Therefore the expected maximum number of applications still awaiting approval in the steady state is given by

$$\begin{aligned}
E[\max\{X^f, X^l\}] &= (1-\rho_f)(1-\rho_l) \sum_{n=0}^{\infty} n(\rho_f \rho_l)^n + \sum_{n=0}^{\infty} n(1-\rho_f)\rho_f^n + \\
&\quad \sum_{n=0}^{\infty} n(1-\rho_l)\rho_l^n - \sum_{n=0}^{\infty} n(1-\rho_f)(\rho_f\rho_l)^n - \sum_{n=0}^{\infty} n(1-\rho_l)(\rho_f\rho_l)^n
\end{aligned}$$

$$\begin{aligned}
&= \sum_{n=0}^{\infty} n(1-\rho_f)\rho_f^n + \sum_{n=0}^{\infty} n(1-\rho_l)\rho_s^n - (1-\rho_f\rho_l) \sum_{n=0}^{\infty} n(\rho_f\rho_l)^n \\
&= \frac{\rho_f}{(1-\rho_f)} + \frac{\rho_l}{(1-\rho_l)} - \frac{\rho_f\rho_l}{1-\rho_f\rho_l}
\end{aligned}$$

Thus

$$ENA\,(SIS) \equiv E[\max\{X^f, X^l\}] = \frac{\rho_f}{(1-\rho_f)} + \frac{\rho_l}{(1-\rho_l)} - \frac{\rho_f\rho_l}{1-\rho_f\rho_l}$$

Under *SES*, once the applications are served by *FG*, they come to *LG* for approval. Note that the the queue faced by *FG* is unaffected by the queue that *LG* faces. We assume $\lambda < \mu_i$, $i = f, l$ for a steady state to exist. The queue faced by *FG* is therefore an isolated queue with arrival rate λ and the approval rate μ_f. The departure from FG's queue forms the arrival to the LG. But this departure forms a Poisson stream with rate λ. Therefore the queue that LG faces is just like a queue where the arrival rate is a Poisson stream with rate λ and the service rate as we assumed is μ_l. If $X^f(t)$ and $X^l(t)$ represent the number of applications in the steady state for FG and LG respectively, then $\pi^*_{n_f,n_l} = \Pr[X^f = n_f \; , \; X^l = n_l] = \Pr[X^f = n_f] \Pr[X^l = n_l]$ and therefore, $\pi^*_{n_f,n_l} = (\rho_f)^{n_f}(1-\rho_f)(\rho_l)^{n_l}(1-\rho_l)$. Given the steady state distribution, we can easily find out the expected number of applications that are awaiting approval in the steady state and which is given by

$$E[X^f + X^l] = \sum_{n_f}\sum_{n_l}(n_f + n_l)(\rho_f)^{n_f}(1-\rho_f)(\rho_l)^{n_l}(1-\rho_l) = \frac{\rho_f}{(1-\rho_f)} + \frac{\rho_l}{(1-\rho_l)}$$

Hence the inefficiency of SES, denoted by $ENA(SES)$, is

$$ENA\,(SES) = \frac{\rho_f}{(1-\rho_f)} + \frac{\rho_l}{(1-\rho_l)}$$

Now we discuss the steady state of SWS. Under SWS, the arrival forms a Poisson process with rate λ and the service rate is also Poisson with a rate $[\frac{1}{2}(\mu_f^{-1} + \mu_l^{-1})]^{-1}$. The steady state distribution is given by

$$\pi^*_n = \Pr[X = n] = (\frac{\lambda}{[\frac{1}{2}(\mu_f^{-1} + \mu_l^{-1})]^{-1}})^n(1 - \frac{\lambda}{[\frac{1}{2}(\mu_f^{-1} + \mu_l^{-1})]^{-1}}) = (\rho^\circ)^n(1-\rho^\circ)$$

where $\rho^\circ = [\frac{1}{2}(\mu_f^{-1} + \mu_l^{-1})]\lambda$. Hence the expected number of applications that are still pending in the steady state is given by $\dfrac{\rho^\circ}{(1-\rho^\circ)}$. We denote the inefficiency under SWS

by $ENA(SWS)$, where

$$ENA(SWS) = \frac{\rho^\circ}{(1-\rho^\circ)} = \frac{\frac{1}{2}(\rho_f + \rho_l)}{1 - \frac{1}{2}(\rho_f + \rho_l)}$$

Given the above results, we prove the following proposition easily.

Proposition 1 *If the inefficiency of an approval system is measured by ENA, the policy SWS is the most efficient, followed by SIS and SES, i.e., $ENA(SWS) < ENA(SIS) < ENA(SES)$.*

Proof. Note that $\frac{\rho}{(1-\rho)}$ is an increasing function of ρ. Hence, clearly

$$\frac{\rho^\circ}{(1-\rho^\circ)} \leqq \max\{\frac{\rho_f}{(1-\rho_f)}, \frac{\rho_l}{(1-\rho_l)}\} \tag{1}$$

But

$$\max\{\frac{\rho_f}{(1-\rho_f)}, \frac{\rho_l}{(1-\rho_l)}\} \leq \frac{\rho_f}{(1-\rho_f)} + \frac{\rho_l}{(1-\rho_l)} - \frac{\rho_f\rho_l}{1-\rho_f\rho_l} \tag{2}$$

as both $\frac{\rho_l}{(1-\rho_l)} - \frac{\rho_f\rho_l}{1-\rho_f\rho_l}$ and $\frac{\rho_f}{(1-\rho_f)} - \frac{\rho_f\rho_l}{1-\rho_f\rho_l} \geqq 0$. Also

$$\frac{\rho_f}{(1-\rho_f)} + \frac{\rho_l}{(1-\rho_l)} - \frac{\rho_f\rho_l}{1-\rho_f\rho_l} < \frac{\rho_f}{(1-\rho_f)} + \frac{\rho_l}{(1-\rho_l)} \tag{3}$$

Thus by combining the above three inequalities, we have

$$\frac{\rho^\circ}{(1-\rho^\circ)} < \frac{\rho_f}{(1-\rho_f)} + \frac{\rho_l}{(1-\rho_l)} - \frac{\rho_f\rho_l}{1-\rho_f\rho_l} < \frac{\rho_f}{(1-\rho_f)} + \frac{\rho_l}{(1-\rho_l)}$$

i.e., $ENA(SWS) < ENA(SIS) < ENA(SES)$. ■

Now let us change the assumption regarding the service rate of a SWS. Instead of $[\frac{1}{2}(\mu_f^{-1} + \mu_l^{-1})]^{-1}$, let the service rate be $\mu_{\min} \equiv \min\{\mu_f, \mu_l\}$, while the arrival rate is fixed at the same level of λ. We denote the inefficiency of SWS associated with the service rate $\mu_{\min}$ by $ENA\,(SWS)^{\min}$.

Proposition 2 *Under ENA, a SWS policy with service rate $\mu_{\min} \equiv \min\{\mu_f, \mu_l\}$ is also more efficient than the SIS and SES, i.e., $ENA\,(SWS)^{\min} < ENA\,(SIS) < ENA\,(SES)$.*

Proof. The expected number of applications awaiting approval in the steady state is $ENA\,(SWS)^{\min} = \frac{\rho_{\min}}{1-\rho_{\min}}$, where $\rho_{\min}$, with slightly abuse of notation, is the ρ associated with $\mu_{\min}$. Note that $\frac{\rho_{\min}}{1-\rho_{\min}} = \max\{\frac{\rho_f}{(1-\rho_f)}, \frac{\rho_l}{(1-\rho_l)}\}$. Thus the rest of the proof follows directly from the proof of proposition 1. ■

This proposition essentially means that even if the service rate of the single window is the minimum of the rates at which the two windows process applications in a multiple window system, single window is still the more efficient approval system. The overall minimum service rate can be achieved by downsizing the total labor force. But this could reduce the employment level of an economy. Alternatively, it is possible to offer the minimum service rate with the same number of employees with a less average workload on each employee. Thus an efficient system through reorganizing the labor force may be achieved even by working less.

2.2 Inefficiency under EWT

The second measure of inefficiency is the expected waiting time of an application in the steady state. Under this measure too, we reach the same conclusion as above that a single window is the most efficient approval system. Let W_k be the service time of the k^{th} application and X be the number of applications in the system just before the arrival of an arbitrary application. Let S_n denote the total time that the n^{th} application spends in the system, where total time is defined by the waiting time plus the service time. We are interested to find out the mean waiting time of an application at the steady state. The limiting distribution of S_n is

$$F_S(t) = \Pr(S \le t) = \lim_{n\to\infty} \Pr(S_n \le t)$$

i.e., $F_S(t)$ measures the probability that the total time that an application spends in the system is less than or equal to t. Also $\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{\infty} S_k = E(S)$ w.p. 1.

We first reproduce a basic result (See Cooper 1981; Klienrock 1975 for a different treatment) that is essential for our study. Note that the exponential service distribution is memoryless and therefore the random variables W_k are independent of each other $\forall k$. As $W_k \sim \mathit{exponential}\,(\mu)\ \forall k$, $E(W_k) = \mu^{-1}$ and $Var(W_k) = \mu^{-2}$. Also note that

$$\Pr(S > t) = \Pr(\sum_{k=1}^{X+1} W_k > t) = \sum_{n=0}^{\infty} \Pr(\sum_{k=1}^{n+1} W_k > t)\Pr(X = n)$$

Now $\sum_{k=1}^{n+1} W_k$ has an Erlang-$(n+1)$ distribution. Therefore,

$$\begin{aligned}\Pr(S &> t,\ t \geqq 0) = \sum_{n=0}^{\infty}\sum_{k=0}^{n} \frac{(\mu t)^k}{k!} e^{-\mu t}(1-\rho)\rho^n \\ &= \sum_{k=0}^{\infty}\sum_{n=k}^{\infty} \frac{(\mu t)^k}{k!} e^{-\mu t}(1-\rho)\rho^n = \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} e^{-\mu t} = e^{-(\mu-\lambda)t}\end{aligned}$$

Thus $S \sim \mathit{exponential}\,(\mu - \lambda)$ and hence $E(S) = (\mu-\lambda)^{-1}$ and $Var(S) = (\mu-\lambda)^{-2}$.

Let g and G represent the probability density function and cumulative distribution function respectively. Also let S^i be associated with $i = f, l$. Now the inefficiency of *SIS*, denoted by $EWT\,(\mathit{SIS})$, can be measured as EU where the density of $U \equiv \max\{S^f, S^l\}$

is given by

$$
\begin{aligned}
g[U &\equiv \max\{S^f, S^l\}] \\
&= g_{S^f}(u)G_{S^l}(u) + g_{_{S^l}}(u)G_{S^f}(u) \\
&= \left(\mu_f - \lambda\right) e^{-\left(\mu_f-\lambda\right)u}(1 - e^{-(\mu_l-\lambda)u}) + (\mu_l - \lambda)\, e^{-(\mu_l-\lambda)u}(1 - e^{-\left(\mu_f-\lambda\right)u}) \\
&= \left(\mu_f - \lambda\right) e^{-\left(\mu_f-\lambda\right)u} + (\mu_l - \lambda)\, e^{-(\mu_l-\lambda)u} - \left(\mu_f + \mu_l - 2\lambda\right) e^{-\left(\mu_f+\mu_l-2\lambda\right)u}
\end{aligned}
$$

Therefore $EWT\,(SIS) \equiv EU = \dfrac{1}{\left(\mu_f - \lambda\right)} + \dfrac{1}{(\mu_l - \lambda)} - \dfrac{1}{\mu_f + \mu_l - 2\lambda}$. The justification for choosing this max operator has been discussed when we calculated $ENA\,(SIS)$ under ENA. Similarly, $EWT\,(SES)$ can be measured as EV where the density of $V \equiv (S^f + S^l)$ is given by

$$
\begin{aligned}
g[V &\equiv S^f + S^l] \\
&= \int_0^v g_{S^f}(v - s^l) g_{_{S^l}}(s^l) ds^l \\
&= \frac{\left(\mu_f - \lambda\right)(\mu_s - \lambda)}{\mu_l - \mu_f}[e^{-\left(\mu_f-\lambda\right)v} - e^{-(\mu_l-\lambda)v}]
\end{aligned}
$$

Therefore, $EWT\,(SES) \equiv EV = \dfrac{1}{\left(\mu_f - \lambda\right)} + \dfrac{1}{(\mu_l - \lambda)}$. Calculating the inefficiency for SWS is straight forward. It can easily be seen that

$$
EWT\,(SWS) = \frac{1}{[\frac{1}{2}(\mu_f^{-1} + \mu_l^{-1})]^{-1} - \lambda}
$$

We are now in a position to compare the inefficiencies as we did under ENA. The next proposition follows easily from the above discussion.

Proposition 3 *If the inefficiency of an approval system is measured by EWT, the policy*

SWS is the most efficient, followed by SIS and SES, i.e., EWT (SWS) < EWT (SIS) < EWT (SES).

Proof. Note that $[\frac{1}{2}(\mu_f^{-1} + \mu_l^{-1})] \leqq \max(\mu_f^{-1}, \mu_l^{-1}) = [\min(\mu_f, \mu_l)]^{-1}$. Therefore

$$\frac{1}{[\frac{1}{2}(\mu_f^{-1} + \mu_l^{-1})]^{-1} - \lambda} \leqq \frac{1}{\min(\mu_f, \mu_l) - \lambda} \tag{4}$$

Since both the terms $\frac{1}{(\mu_l - \lambda)} - \frac{1}{\mu_f + \mu_l - 2\lambda}$ and $\frac{1}{(\mu_f - \lambda)} - \frac{1}{\mu_f + \mu_l - 2\lambda} > 0$ we have

$$\frac{1}{\min(\mu_f, \mu_l) - \lambda} < \frac{1}{(\mu_f - \lambda)} + \frac{1}{(\mu_l - \lambda)} - \frac{1}{\mu_f + \mu_l - 2\lambda} \tag{5}$$

Clearly

$$\frac{1}{(\mu_f - \lambda)} + \frac{1}{(\mu_l - \lambda)} - \frac{1}{\mu_f + \mu_l - 2\lambda} < \frac{1}{(\mu_f - \lambda)} + \frac{1}{(\mu_l - \lambda)} \tag{6}$$

Thus combining the above three inequalities, we have

$$\frac{1}{[\frac{1}{2}(\mu_f^{-1} + \mu_l^{-1})]^{-1} - \lambda} < \frac{1}{(\mu_f - \lambda)} + \frac{1}{(\mu_l - \lambda)} - \frac{1}{\mu_f + \mu_l - 2\lambda} < \frac{1}{(\mu_f - \lambda)} + \frac{1}{(\mu_l - \lambda)}$$

i.e., $EWT\,(SWS) < EWT\,(SIS) < EWT\,(SES)$. ■

Thus under the measure EWT too, SWS performs better than the SIS and SES. We now proceed with presenting the following result which is the corresponding result of proposition 2 under the measure EWT. Here we denote the inefficiency of SWS associated with the service rate $\mu_{\min}$ by $EWT\,(SWS)^{\min}$.

Proposition 4 *Under EWT, a SWS policy with service rate* $\mu_{\min} \equiv \min\{\mu_f, \mu_l\}$ *is also more efficient than the SIS and SES, i.e.,* $EWT\,(SWS)^{\min} < EWT\,(SIS) < EWT\,(SES)$.

Proof. Note that $EWT\,(SWS)^{\min} = \dfrac{1}{\min(\mu_f, \mu_l) - \lambda}$. The rest of the proof follows directly from the proof of proposition 3. ■

2.3 Some facts and figures

In this subsection we provide some results regarding the inefficiencies of different clearance systems at the steady state for different values of λ, μ_f and μ_l. The choice of the values of the parameters is arbitrary, subject to the restrictions required by the stability conditions. We place each of the figures under ENA and EWT side by side. We have used MATLAB to generate the figures. For the first four figures (Figures 1 to 4), we fix μ_f and μ_l at 180 and 200 respectively, and increase λ from 40 to 150. Figure 1 shows that given that the values of μ_f and μ_l are fixed, the inefficiency measured under ENA increases with an increase in the the arrival rate of applications. Figure 2 is the corresponding version for the measure EWT. These results are quite expected. However, it is interesting to note that the inefficiency gap between the single window and other systems also increases with λ, given that values of μ_f and μ_l are fixed. Figures 3 and 4 plot the gap between the inefficiencies of the three systems for ENA and EWT respectively. We denote $JK = ENA(SJS) - ENA(SKS)$ or $JK = EWT(SJS) - EWT(SKS)$ whichever is applicable, where $J, K = I, E, W$. As an example, IW in Figure 3 denotes the gap between the inefficiencies of the systems SIS and SWS, when inefficiency is measured by ENA. Similarly, IW in Figure 4 denotes the gap between the inefficiencies of the systems SIS and SWS, when inefficiency is measured by EWT. We see that under both ENA and EWT, IW and EW increase with λ. This result tells us that in a growing economy, where an increase in economic activity leads to more people applying for passports, investors applying for approvals for new projects, potential exporters applying for a license etc, SWS performs increasingly better.

For the remaining four figures (Figures 5 to 8), we fix μ_l and λ at 200 and 40 respectively, and increase μ_f from 150 to 1200. We could alternatively fix μ_f and λ and vary μ_l. Figure 5 shows how the inefficiency of the clearance systems under ENA changes as the service rate of FG improves, keeping the arrival rate and service rate of LG fixed. Figure 6 is the corresponding version for the measure EWT. These results are also quite expected. Figure 7 shows that the both EW and IW decrease as μ_f increases. Figure 8 shows that after initially decreasing, IW increases after a particular value of μ_f. The results that figures 7 and 8 depict are quite consistent. The shape of the curves in the last two figures are not unique in the sense that their shapes vary with the relative value of μ_f and μ_l as well as λ. As an example, if we fix λ and μ_l at 5 and 6 respectively and vary μ_f from 6 to 14, we observe that both IW and EW first decrease and then start increasing after a particular value of μ_f under ENA as well as EWT. If we keep on increasing μ_f to a very large number, keeping μ_l and λ constant, we see that the gap in inefficiencies between SWS and the others converges to a particular value though unnecessary increment of μ does not make much sense in our discussion.

[Figures 1-8 here]

3 The Structure of Single Window

In the previous section we have established that an approval system where there is a single window with a single server is more efficient than multiple window approval systems. However, the agency which acts as a single window could place its resources in front of its queue of applications in many ways. One could have a single window approval system where there are multiple servers. Note that in our paper the concept of a single window is slightly different than that defined by Athreya and Majumdar (2005). Suppose the applicant has to submit one combined application to a single agency. However, for the ease of the applicants the agency has two sets of people (servers) who have the authority to approve an application. The applicants can form two queues in front of the two servers. Unlike Athreya and Majumdar (2005), in our paper we treat this type of an approval process as a SWS. The justification for this is that in spite of there being two queues, the applications have to gain approval of one agency only. The *single window single server* can be regarded as the extreme version of the service provided by a SWS where the application is 'one single complex license' that takes care of everything (see for e.g. L-M-R 2007), and people with different expertise come together to form only one server. Now the question arises whether a single window with a single server is superior to the other forms of single window. In this section we answer this question in the affirmative.

Let $SWS(k \mid l)$ denote a single window system where k and l represent the number of queues and the number of servers respectively. The three possible structures of the single window that we consider under the $M/M/(.)$ framework are as follows:

$SWS(2 \mid 2)$: In this system there are two separate queues each having arrival rate $\frac{\lambda}{2}$ and two separate servers each having service rate of μ. For simplicity, we assume that the service rates are same for all the servers of the single window. The arrival processes of the two queues are assumed to be independent. The service processes of the two servers are also independent of each other. Furthermore, the arrival process of applications is independent of the service process for each of the queues.

$SWS(1 \mid 2)$: In this system, instead of having two, there is only one queue with a combined arrival rate of application λ and two counters each having service rate of μ. The arrival process of applications is assumed to be independent of the service process and the service processes of the two servers are also independent of each other.

$SWS(1 \mid 1)$: Under this system there is only one queue with a combined arrival rate λ and only one server with the combined service rate $\theta\mu$. This can be termed as a *single window single server* system. Here too the arrival process of applications is independent of the service process.

In the following subsection we compare the inefficiencies of $SWS(2 \mid 2)$, $SWS(1 \mid 2)$ and $SWS(1 \mid 1)$ under the measures ENA and EWT.

3.1 Inefficiency under ENA and EWT

Under ENA, the expected queue size at the steady state for $SWS(2 \mid 2)$ can be measured as $ENA(SWS(2 \mid 2)) = 2EX = \frac{2\rho}{2-\rho}$ while $ENA(SWS(1 \mid 1)) = \frac{\rho}{\theta-\rho}$ where $\rho = \frac{\lambda}{\mu}$. These two results can be easily arrived at from the discussion under section 2 where each and every queue is a simple $M/M/1$ process. $SWS(1 \mid 2)$ is a $M/M/2$ system and it can be seen that $ENA(SWS(1 \mid 2)) = EX = \frac{4\rho}{4-\rho^2}$ (See Gross and Harris 1985, pp 88; Durrett 1999, pp 179).

Note that $EWT(SWS(2 \mid 2)) = \frac{2\rho}{\lambda(2-\rho)}$ as the length of each of the queues is $\frac{\rho}{(2-\rho)}$ while the arrival rate is $\frac{\lambda}{2}$. It is straight forward to show that $EWT(SWS(1 \mid 2)) = \frac{4\rho}{\lambda(4-\rho^2)}$ and $EWT(SWS(1 \mid 1)) = \frac{\rho}{\lambda(\theta-\rho)}$. Hence we have the following proposition.

Proposition 5 *There exists a θ^c such that if $\theta > \theta^c$, under both the measures ENA and EWT, $SWS(1 \mid 1)$ is the most efficient, followed by $SWS(1 \mid 2)$ and ($SWS(2 \mid 2)$).*

Proof. Note that for stability we require $\theta > \rho$. Also note that $ENA(SWS(1 \mid 2)) < ENA(SWS(2 \mid 2))$ and $EWT(SWS(1 \mid 2)) < EWT(SWS(2 \mid 2))$ hold. Clearly $ENA(SWS(1 \mid 1)) < ENA(SWS(1 \mid 2))$ and $EWT(SWS(1 \mid 1)) < EWT(SWS(1 \mid 2))$ hold when $\theta > 1 + \rho - (\frac{\rho}{2})^2 \equiv \theta^c$. Note that the stability condition is also satisfied as $1 + \rho - (\frac{\rho}{2})^2 > \rho$. Hence proved. ■

It can easily be seen that if the service rate of $SWS(1 \mid 1)$ is 2μ (i.e., $\theta = 2$), proposition 5 clearly holds. Note that a single window with a single server of service rate 2μ always has departures at a rate 2μ whereas two servers each having service rate μ sometimes has a departure rate less than 2μ. Though finding out an exact rate of service of the combined labor force under $SWS(1 \mid 1)$ is under the area of empirical investigation, feasibility of having a $SWS(1 \mid 1)$ with a service rate of 2μ cannot be a question. For example, suppose the processing time of an application by a set of employees under the $SWS(1 \mid 1)$ is exponentially distributed with a mean time say one month i.e., the service rate of application is one per month. It is reasonable to argue that two identical sets of employees working under a single umbrella will be able to serve an application in half of a month, i.e., jointly they can serve two applications per month, even if there is no additional gain from having two identical groups work together. In fact we do not even need a service rate of 2μ for the combined server all the time. Note that the threshold value of θ, θ^c is an increasing function of ρ and when $\rho \to 0$, the required combined rate of service is simply greater than μ. Thus, the threshold value of θ can well be achieved.

We now extend our analysis in a more general setup, namely $M/G/(.)$, i.e., the Poisson arrival and the service processes follow general distribution. We show that it is possible to compare the above three inefficiencies in the general setup too. The same framework has been considered by Athreya and Majumdar (2005). We could consider an even more general framework $GI/G/(.)$, i.e., instead of Poisson arrival, it is a general input arrival process. However, this framework with many approximation results impose unnecessary restrictions on our analysis.

3.2 A more general framework

Here we consider a $M/G/(.)$ framework. Let X_n be the number of applications in a queue at the moment just after the approval of application n finishes. With Poisson arrivals with rate λ and the service times are i.i.d. with distribution G, $\{X_n\}_{n=0}^{\infty}$ is a Markov Chain with the state space $J = \{0, 1, 2, ...\}$ that satisfies

$$X_{n+1} = \begin{cases} X_n - 1 + \varepsilon_{n+1} & \text{if } X_n \geqq 1 \\ \varepsilon_{n+1} & \text{if } X_n = 0 \end{cases}$$

where $\{\varepsilon_n\}_{n \geqslant 1}$ are i.i.d. random variables with ε_n represents the number of arrivals during n^{th} customer's service time. The first order transition probabilities of $\{X_n\}$ are given by $p_{0j} = p_{1j} = p_j = \int_0^{\infty} (\lambda t)^j [\frac{\exp(-\lambda t)}{j!}] dG(t)$, $j > 0$. For $i > 1$, $p_{ij} = p_{j-i+1}$.

It can be shown that there is a unique steady state distribution $\{\pi_j\}$ and X_n converges in distribution to a random variable X where $\Pr(X = j) = \pi_j$. In the steady state, one can show that $EX = \frac{\lambda^2 E(S')^2}{2(1-\rho)} + \rho$ provided $E(S')^2 < \infty$ where S' is the random variable that represents the service time of the system with distribution G and $\rho = \lambda ES'$. On the other hand we are not aware of any particular result for the expected queue length and the expected waiting time for a $M/G/k$ system. There are few approximations available for a $M/G/k$ system. We use the approximation of mean waiting time *in the queue,* EW_q, due to Nozaki and Ross (1978). On the accuracy of using the approximation by Nozaki and Ross, we argue that it is consistent with the Kingman's (Kingman 1970) lower and upper bound on W_q for $GI/G/k$ when adapted to $M/G/k$. Thus by Little's law, the expected queue length at the stochastic steady state is $EX \cong \lambda(\frac{\lambda^k E(S')^2 (ES')^{k-1}}{2(k-1)!(k-\lambda ES)^2(\sum_{n=0}^{k-1} \frac{(\lambda ES)^n}{n!} + \frac{(\lambda ES)^k}{(k-1)!(k-\lambda ES)})} + ES')$, $E(S')^2 < \infty$. Using the approximation by Nozaki and Ross for $k = 2$, we have

$EX \cong \lambda(\frac{\lambda^2 E(S')^2 ES'}{2(2-\lambda ES')^2(1+\lambda ES' + \frac{(\lambda ES')^2}{(2-\lambda ES')})} + ES').$

Here $SWS(2 \mid 2)$ consists of two $M/G/1$ queues with each queue having the characteristics $(\frac{\lambda}{2}, S', \frac{1}{ES'})$. That is, in $SWS(2 \mid 2)$, the arrival rate is $\frac{\lambda}{2}$ for each of the two queues and the mean service time is ES' for each of the two servers. The second system namely $SWS(1 \mid 2)$ is a $M/G/2$ queue which has the characteristics $(\lambda, S', \frac{1}{ES'})$. That is, under $SWS(1 \mid 2)$, there is only one queue with a combined arrival rate of λ but there are two servers each having the expected service time ES'. The third system $SWS(1 \mid 1)$ is a $M/G/1$ characterized by $(\lambda, S, \frac{1}{ES})$ and represents the *single window single server* system. In $SWS(1 \mid 1)$, there is only one queue simply like the $SWS(1 \mid 2)$ system with a combined arrival rate of λ and there is only one server with a combined service rate of $\frac{1}{ES}$, that is, we assume that the expected service time of this $SWS(1 \mid 1)$ system is ES. By construction, the three systems described above satisfy our notion of single window.

3.2.1 Inefficiency under ENA and EWT

We start with the measure ENA. We know that

$$ENA(SWS(2 \mid 2)) = 2EX = \frac{\lambda^2 E(S')^2}{2(2-\lambda ES')} + \lambda ES'$$

We also have

$$ENA(SWS(1 \mid 2)) = EX \cong \lambda(\frac{\lambda^2 E(S')^2 ES'}{2(2-\lambda ES')^2(1+\lambda ES' + \frac{(\lambda ES')^2}{(2-\lambda ES')})} + ES')$$

and

$$ENA(SWS(1 \mid 1)) = \frac{\lambda^2 E(S)^2}{2(1-\lambda ES)} + \lambda ES$$

Let us assume that S' has the same first two moments as $\theta' S$ where $\theta' \equiv \frac{\mu}{\mu'}$.

Proposition 6 *Under* ENA, $\exists\, \overline{\theta} \ni \forall \theta' \in [\overline{\theta}, \frac{2}{\lambda ES})$, $ENA(SWS(1 \mid 1)) < ENA(SWS(1 \mid 2)) < ENA(SWS(2 \mid 2))$ *holds.*

Proof. Note that

$$\frac{ENA(SWS(2 \mid 2))}{ENA(SWS(1 \mid 2))} \equiv \gamma(\theta') = \frac{\dfrac{\lambda^2 E(S')^2}{2(2-\lambda ES')} + \lambda ES'}{\lambda(\dfrac{\lambda^2 E(S')^2 ES'}{2(2-\lambda ES')^2(1+\lambda ES' + \frac{(\lambda ES')^2}{(2-\lambda ES')})} + ES')}$$

$$= \frac{\dfrac{\lambda E(S')^2}{2(2-\lambda ES')} + ES'}{\dfrac{\lambda^2 E(S')^2 ES'}{2(2-\lambda ES')(2+\lambda ES')} + ES'}$$

Hence to show that $\gamma(\theta') > 1$, it is sufficient to show that $\dfrac{\dfrac{\lambda E(S')^2}{2(2-\lambda ES')}}{\dfrac{\lambda^2 E(S')^2 ES'}{2(2-\lambda ES')(2+\lambda ES')}} > 1$.

Note that $\dfrac{\dfrac{\lambda E(S')^2}{2(2-\lambda ES')}}{\dfrac{\lambda^2 E(S')^2 ES'}{2(2-\lambda ES')(2+\lambda ES')}} > 1 \Rightarrow \frac{2+\lambda ES'}{\lambda ES'} = 1 + \frac{2}{\lambda ES'} > 1$. Obviously, the inequality is always true. Now we compare $ENA(SWS(1 \mid 2))$ and $ENA(SWS(1 \mid 1))$. We have

$$\frac{ENA(SWS(1 \mid 1))}{ENA(SWS(1 \mid 2))} \equiv \varphi(\theta') = \frac{\dfrac{\lambda^2 E(S)^2}{2(1-\lambda ES)} + \lambda ES}{\lambda(\dfrac{\lambda^2 E(S')^2 ES'}{2(2-\lambda ES')^2(1+\lambda ES' + \frac{(\lambda ES')^2}{(2-\lambda ES')})} + ES')}$$

$$= \frac{\dfrac{\lambda^2 E(S)^2}{2(1-\lambda ES)} + \lambda ES}{\lambda(\dfrac{\theta'^3 \lambda^2 E(S)^2 ES}{2(2-\theta'\lambda ES)(2+\theta'\lambda ES')} + \theta' ES)}$$

Note that for stability we require $\lambda ES' < 2 \Rightarrow \lambda\theta' ES < 2 \Rightarrow \theta' < \frac{2}{\lambda ES}$. We also observe that $\varphi(\theta')$ is monotonically decreasing, with $\lim_{\theta' \to 0} \varphi(\theta') = \infty$ and $\lim_{\theta' \to \frac{2}{\lambda ES}} \varphi(\theta') = 0$. Therefore, $\exists\ \overline{\theta} \ni \varphi(\theta') < 1\ \forall \theta' \in [\overline{\theta}, \frac{2}{\lambda ES})$. Given $\gamma(\theta') > 1$, the last result is enough to prove our proposition. However we will discuss all the three single window systems separately, specify the domain of θ' for each of the comparisons where the desired result holds and draw the conclusion thereafter. Thus we compare $ENA(SWS(2 \mid 2))$ and $ENA(SWS(1 \mid 1))$ in a similar way. We observe that

$$\begin{aligned}\frac{ENA(SWS(1 \mid 1))}{ENA(SWS(2 \mid 2))} &\equiv \phi(\theta') = \frac{\dfrac{\lambda^2 E(S)^2}{2(1-\lambda ES)} + \lambda ES}{\dfrac{\lambda^2 E(S')^2}{2(2-\lambda ES')} + \lambda ES'} \\ &= \frac{\dfrac{\lambda^2 E(S)^2}{2(1-\lambda ES)} + \lambda ES}{\dfrac{\theta'^2 \lambda^2 E(S)^2}{2(2-\theta'\lambda ES)} + \theta'\lambda ES}\end{aligned}$$

Here, for stability we require $\frac{\lambda}{2} ES' < 1 \Rightarrow \lambda\theta' ES < 2 \Rightarrow \theta' < \frac{2}{\lambda ES}$ and hence the same restriction on θ' works for both $\varphi(\theta')$ and $\phi(\theta')$. Clearly $\phi(\theta')$ is also monotonically decreasing, with $\lim_{\theta' \to 0} \phi(\theta') = \infty$ and $\lim_{\theta' \to \frac{2}{\lambda ES}} \phi(\theta') = 0$. Therefore, $\exists\ \widetilde{\theta} \ni \phi(\widetilde{\theta}) < 1$ $\forall \theta' \in [\widetilde{\theta}, \frac{2}{\lambda ES})$. Now we combine the above three comparisons of inefficiencies together to prove our proposition. Given that $\gamma(\theta') > 1$ holds, we must have $\phi(\theta') < \varphi(\theta')$ $\forall \theta' \in [0, \frac{2}{\lambda ES})$. Therefore, $\phi(\theta') \mid_{\theta' = \overline{\theta}} << 1$, i.e., $\widetilde{\theta} < \overline{\theta}$. Also note that $\overline{\theta} > 1$. Thus, $ENA(SWS(1 \mid 1)) < ENA(SWS(1 \mid 2)) < ENA(SWS(2 \mid 2))$ holds for $\theta' \in [\widetilde{\theta}, \frac{2}{\lambda ES}) \cap [\overline{\theta}, \frac{2}{\lambda ES}) = [\overline{\theta}, \frac{2}{\lambda ES})$. ■

The values of $E(S)^2$ and $E(S')^2$ also play a crucial role here. Note that $Var(S') = E(S')^2 - [E(S')]^2 = \theta'^2\ Var(S) \Rightarrow \theta'^2 = \frac{E(S')^2}{E(S)^2} = \frac{Var(S')}{Var(S)}$. As $\varphi'(\theta') < 0$, the value of θ' reaches $\overline{\theta}$ fast if $\frac{Var(S')}{Var(S)}$ is high. To be more elaborate, instead of assuming that S' and $\theta' S$

have the same first two moments, if we assume that only the first moment is the same, we can check that if we set a particular value for $\overline{\theta}$, a high $E(S)^2$ compared to $E(S')^2$ may result in $\frac{ENA(SWS(1|1))}{ENA(SWS(1|2))} > 1$. If we keep on lowering the value of $E(S)^2$, we arrive at a point after which $\frac{ENA(SWS(1|1))}{ENA(SWS(1|2))} < 1$ holds. Thus higher the variance of $SWS(1 \mid 1)$, higher is the combined rate of service required for $SWS(1 \mid 1)$ to have $\frac{ENA(SWS(1|1))}{ENA(SWS(1|2))} < 1$. Hence in this general case, for a given θ, a $SWS(1 \mid 1)$ may perform better only when the second order moment of the service time is not very high compared to the other SWS systems. So unlike the comparisons in the $M/M/(.)$ framework of section 3.1, we cannot draw an unambiguous comparison on the entire domain of θ' independent of the value of $E(S)^2$ and $E(S')^2$. The required relationship between the second order moments of the two systems can be derived from the inequality itself. However it does not seem much reasonable that when the labor force is combined to form a $SWS(1 \mid 1)$ system, it should have a very high second moment.

It can easily be shown that under the measure EWT, the same relationship holds in the above specified domain of θ'. The total waiting time of an application in the system is given by $EW = EW_q + ES$. Thus $EWT(SWS(2 \mid 2)) = \dfrac{\lambda E(S')^2}{(2 - \lambda ES')} + ES'$,

$$EWT(SWS(1 \mid 2)) = EW_q + ES \cong \frac{\lambda^2 E(S')^2 ES'}{2(2 - \lambda ES')^2(1 + \lambda ES' + \frac{(\lambda ES')^2}{(2-\lambda ES')})} + ES'$$

whereas $EWT(SWS(1 \mid 1)) = \dfrac{\lambda E(S)^2}{2(1 - \lambda ES)} + ES$.

Proposition 7 *Under* EWT, $\exists\, \overline{\theta} \ni \forall \theta' \in [\overline{\theta}, \frac{2}{\lambda ES})$, $EWT(SWS(1 \mid 1)) < EWT(SWS(1 \mid 2)) < EWT(SWS(2 \mid 2))$ *holds.*

Proof. Note that under EWT, $\dfrac{EWT(SWS(2 \mid 2))}{EWT(SWS(1 \mid 2))} \equiv \xi(\theta') = 2(1 + \frac{2}{\lambda ES'}) = 2\gamma(\theta')$ and hence $\xi(\theta') > 1$. It easily follows from the above that

$$\frac{EWT(SWS(1 \mid 1))}{EWT(SWS(1 \mid 2))} \equiv \nu(\theta') = \frac{\frac{\lambda E(S)^2}{2(1-\lambda ES)} + ES}{\frac{\lambda^2 E(S')^2 ES'}{2(2-\lambda ES')^2(1+\lambda ES' + \frac{(\lambda ES')^2}{(2-\lambda ES')})} + ES'} = \varphi(\theta') \text{ and}$$

$$\frac{EWT(SWS(1 \mid 1))}{EWT(SWS(2 \mid 2))} \equiv \psi(\theta') = \frac{\frac{\lambda E(S)^2}{2(1-\lambda ES)} + ES}{\frac{\lambda E(S')^2}{2(2-\lambda ES')} + ES'} = \frac{\phi(\theta')}{2}$$

Clearly for stability the same restriction $\theta' < \frac{2}{\lambda ES}$ is required here . Also note that if $\theta' \in [\overline{\theta}, \frac{2}{\lambda ES})$, $\nu(\theta') < 1$ holds. But $\psi(\theta') < 1$ for $\theta' \in [\widehat{\theta}, \frac{2}{\lambda ES})$ where $\widehat{\theta} < \widetilde{\theta}$. Therefore $EWT(SWS(1 \mid 1)) < EWT(SWS(1 \mid 2)) < EWT(SWS(2 \mid 2))$ when $\theta' \in [\widehat{\theta}, \frac{2}{\lambda ES}) \cap [\overline{\theta}, \frac{2}{\lambda ES}) = [\overline{\theta}, \frac{2}{\lambda ES})$. Note that $ENA(SWS(1 \mid 1)) < ENA(SWS(1 \mid 2)) < ENA(SWS(2 \mid 2))$ holds for the same domain of θ'. Hence the proof. ■

We can have the same discussion about the role of second order moment of S and S' for the measure EWT.

4 Conclusion

In this paper, we have shown that the single window approval system is the most efficient among the other prevailing clearance systems. We have also established that a re-organization of labor force towards making it a single window system can be effected without a change in government employment. Furthermore, we shed some light on the extreme version of the single window policy and show that a single window single server approval process can be more efficient than a single window with many servers.

References

[1] Athreya, Krishna B., Majumdar, M.: A note on closing the windows: Some lessons from queuing theory. International Journal of Economic Theory **1**, 73-81 (2005)

[2] Bardhan, P.: Corruption and development: a review of issues. Journal of Economic Literature **35**, 1326-46 (1997)

[3] Bhagwati, J.: India in the International Economy. Institute of Public Enterprise. Institute of Public Enterprise, Hyderabad (1973)

[4] Cooper, Robert B.: Introduction to Queueing theory. Macmillan, Second Edition, Elsevier North Holland, New York (1981)

[5] Durrett, R.: Essentials of Stochastic Processes. Springer, New York (1999)

[6] Gross, D., Harris, Carl M.: Fundamentals of Queueing Theory. Second Edition, John Wiley, New York (1985)

[7] Kingman, J.F.C.: Inequalities in the theory of queues. Journal of the Royal Statistical Society. Series B **32**, 102-110 (1970)

[8] Klienrock, L.: Queueing Systems, Volume I: Theory. John Wiley, New York (1975)

[9] Lambart-Mogiliansky, A., Majumdar, M., and Radner, R.: Strategic analysis of petty corruption: Entrepreneurs and bureaucrats. Journal of Development Economics **83**, 351-367 (2007)

[10] Lambart-Mogiliansky, A., Majumdar, M., and Radner, R.: Petty Corruption: A Game Theoretic Approach. International Journal of Economic Theory **4**, 273-297 (2008)

[11] Lui, F.T. : An Equilibrium Queuing Model of Bribery. Journal of Political Economy **93**, 760-81 (1985)

[12] Nozaki, S., and Ross, S.: Approximations in finite capacity multiserver queues with poisson arrivals. Journal of Applied Probability **15**, 826-34 (1978)

[13] Roy, J. Feeling good, but vision still needed, Op-Ed section. The Financial Express, Mumbai. available http://www.financialexpress.com/news/Feeling-Good-But-Vision-Still-Needed/94642/ (2003). Accessed 28 March 2008

[14] Schleifer, A., and Vishney, R.W,: Corruption. Quarterly Journal of Economics **108**, 599-612 (1993)

[15] United Nations Center for Trade Facilitation (UN/CEFACT): Recommendation and Guidelines on Establishing a Single Window, Recommendation No. 33. New York and Geneva. available http://www.unece.org/cefact/recommendations/rec33/rec33_trd352e.pdf (2005). Accessed 28 March 2008

[16] World Bank. Doing Business in 2008 Report. available http://www.doingbusiness.org/documents/FullReport/2008/DB08_Full_Report.pdf (2008). Accessed 28 March 2008

[17] World Bank. Enterprise Surveys. available http://www.enterprisesurveys.org/ (2008). Accessed 28 March 2008

[18] Yoo, H. S.: Petty Corruption. Economic Theory **37**, 267-280 (2008)

Figures

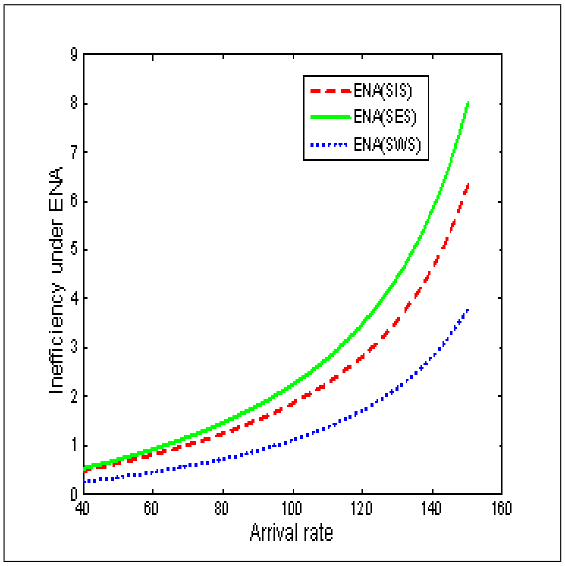

Figure 1: The relationship between inefficiency and λ under ENA

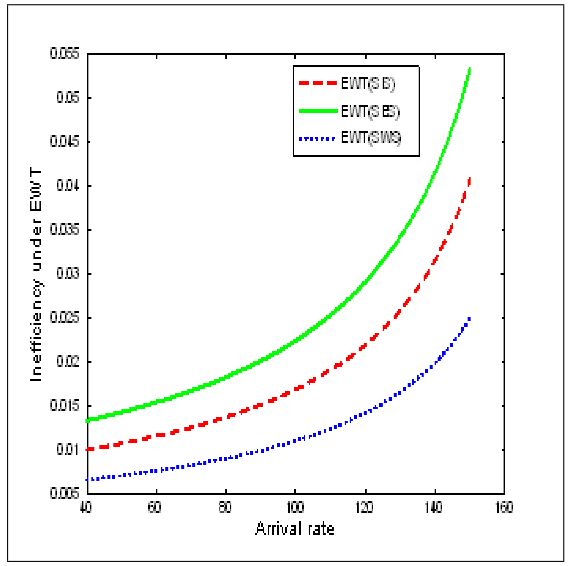

Figure 2: The relationship between inefficiency and λ under EWT

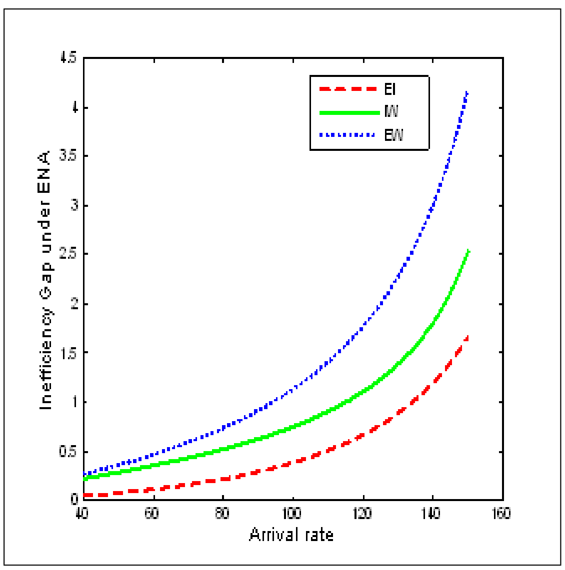

Figure 3: The relationship between gap in the inefficiencies and λ under ENA

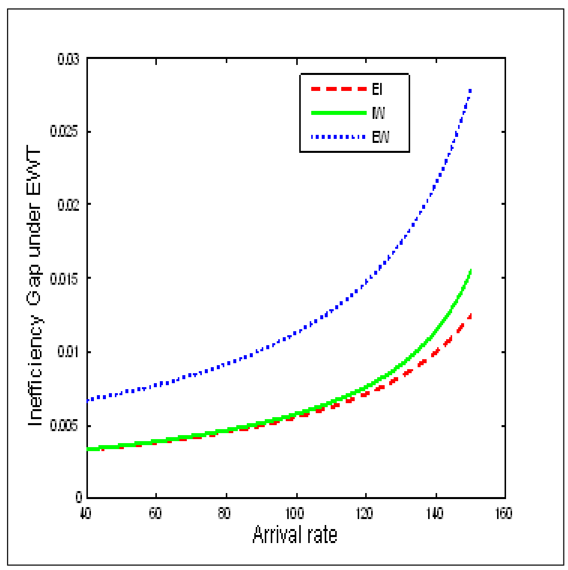

Figure 4: The relationship between the gap in inefficiencies and λ under EWT

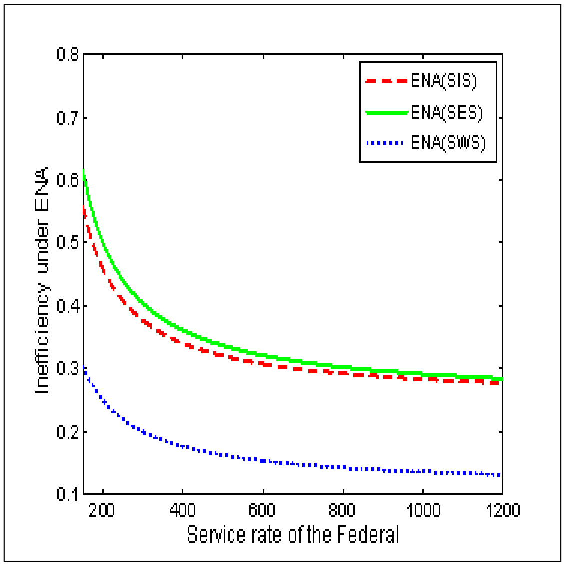

Figure 5: The relationship between inefficiency and μ_f under ENA

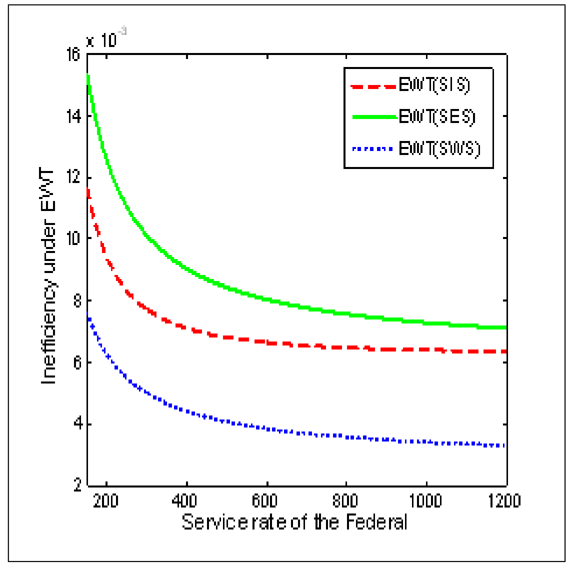

Figure 6: The relationship between inefficiency and μ_f under EWT

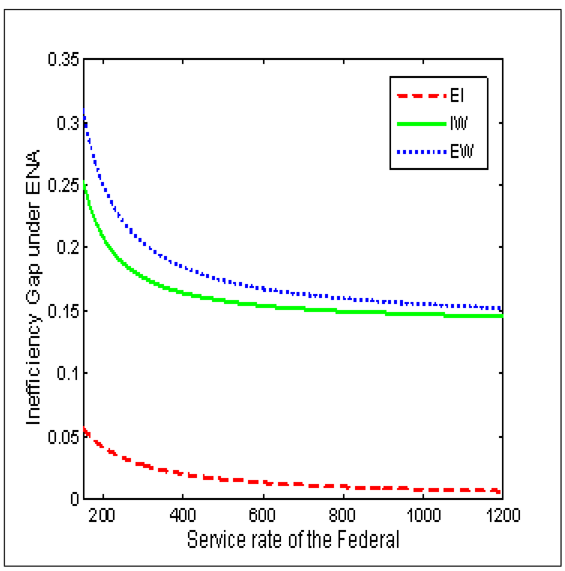

Figure 7: The relationship between the gap in inefficiencies and μ_f under ENA

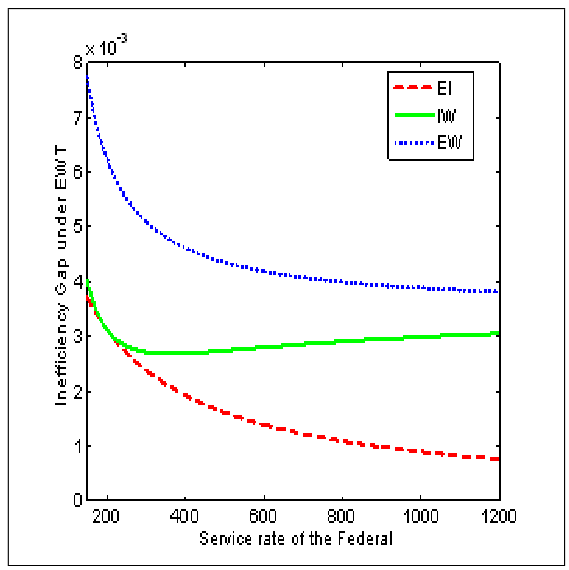

Figure 8: The relationship between the gap in inefficiencies and μ_f under EWT

CHAPTER 3. GOVERNMENT INTERVENTION IN EDUCATION IN THE PRESENCE OF CONSUMPTION EXTERNALITIES

1 Introduction

Governments around the world routinely intervene in the arena of education, and at all levels. While government involvement in primary and secondary education is almost universal, its presence in higher education (especially, in the form of education loans and subsidies) is also fairly common. Heckman (2000), for example, estimates that around 80% of the direct cost at major public universities in the United States is subsidized. The standard rationale for such intervention is the alleged external benefits from education: "existing graduates [will have] more graduates to talk to" [Layard, 1980]. In other words, if there are spillovers from education, then the argument is that those producing it should not bear the full cost of it (for if they were forced to, they would "undereducate" themselves). Such a human-capital-externality argument features prominently in some of the most influential work in growth theory (such as Lucas, 1988) from the last three decades. From a policy-making perspective, the human capital externality and its size is important because it determines the extent to which education and job training should be subsidized (Heckman and Klenow, 1998 and Heckman, 2000).

Yet, the empirical backing for such an externality is not as strong as one would think. Studies involving the estimation of a Mincerian wage equation using micro-level U.S. data, show that the private return to schooling ranges from 4% to 16%, with a consensus estimate around 10% (for a survey, see Card, 1999). Using U.S. state-level data to estimate a Mincerian wage equation, Yamarik (2008) finds that the social return to U.S. schooling is of the order of 9-16%. Turner et. al. (2007) have also used a Mincer model on aggregate-level data and estimated the social return to education in the U.S.

to be 12%-15%. It is the closeness of the estimates of the social return to the private return which suggests that U.S. schooling generates little external return.[1] These sorts of observations help motivate a fairly broad question: how can economists rationalize such ubiquity and importance of government intervention in education when direct spillovers of education are supposedly mild?[2]

In this paper, I narrow the focus of this question, somewhat, by raising it within the context of a neoclassical overlapping generations (OG) model in which *no* intra and intergenerational human capital externalities are present and capital markets are perfect. I argue that externalities in consumption[3] may help make the case for public involvement in education, even in this case. Specifically, I allow for the possibility of intra and intergenerational consumption externalities; an agent faces a consumption floor determined by every other agent alive at that point in her life cycle. This is a natural extension of the keeping-up-with-the-Joneses phenomenon in that living generations compare their consumption to both their own and that of other generations alive. Since agents' saving behavior get affected if they are preoccupied with "keeping up", the allocation of resources to production of human capital in the competitive economy may not be optimal.

The model economy is similar in many respects to the one studied in Boldrin and Montes (2005) and Docquier et al (2007). Agents live for three periods in a OG economy. In the first period of their lives (i.e., when they are young), they do not consume but borrow from a perfect credit market and invest in their human capital. When middle

[1] For more on the external return to education, see Acemoglu and Angrist (2000), Rudd (2000), Ciccone and Peri (2006), and Lange and Topel (2006).

[2] Other popular justifications for strong government involvement in education include the following: a) credit market imperfections: see Becker (1975), Schultz (1961) and Kodde and Ritzen (1985); b) production of social capital (see Putnam et. al. (1993), Durlauf and Fafchamps (2004)); c) creating a civic society of knowledgeable voters (for e.g., Dee, (2004)).

[3] The idea of a consumption externality has long been studied and documented, and involves the notion that people care not just about their own absolute level of consumption but also about how it compares to those around them. Models with interdependent preferences have become immensely popular in many areas of economics and finance. For a recent review, see Abel (2005), Alonso-Carrera et. al (2008), Barnett and Bhattacharya, (2008), among others.

aged, they supply their labor inelastically, produce the final good, repay the education loan, and save in the form of physical capital. When old, they retire and consume the return on their saving. Utility is derived not from the absolute level of consumption but from the difference between the absolute and a reference level of consumption. The reference level of consumption at every time point is an weighted arithmetic average of per capita consumption of the two living generations. A middle-aged agent compares her consumption to that up her peers and also to the consumption of her parents' generation. Similarly, an agent when old, keeps up with the consumption of her own peers and the consumption of the middle-aged generation.

I work out the conditions under which the market economy would over or under accumulate human capital relative to what a benevolent social planner would have deemed correct. There is a technical novelty that deserves mention. While the first-order optimality conditions for the competitive market solution involve the same agent over the two periods of his life, the optimality conditions for the planner involve two different agents (an old and a young agent) alive at the *same* date. If one focused only on steady states, this would not pose any problems. However, since I solve a infinite, dynamic-planning problem with declining weights on generations to come, I have to find a way to make the market and the planning solutions directly comparable if I am to compare the allocations. I do that by devising the notion of a 'laissez-faire supported' social weight; by construction, at this specific planning discount rate, the market and planning optimality *conditions* coincide. However, in the presence of externality in the economy, the market solutions and planning *allocations* at this 'laissez-faire supported' social weight may differ. I begin the discussion by studying the effect of consumption externality at the 'laissez-faire supported' social weight, since at this particular social weight, if there is no externality in the economy, planners optimality *conditions* as well as the *allocations* coincide with the laissez-faire. The comparisons then proceed by contrasting competitive allocations with those preferred by utilitarian planners with discount rates different from *this* weight.

The presence of the consumption externalities ensures that the decentralized market arrangement is typically not Pareto efficient. In fact, the competitive solution (assuming all markets are competitive) deviates from that of the utilitarian social planners' favored allocation in two dimensions: the market economy may over or under accumulate *both* physical and human capital. From Diamond (1965), it is well known that a competitive economy may overaccumulate physical capital if the real return to capital is "too low" (below unity in a model sans population growth). In contrast, I find that both over and under accumulation of physical capital is possible even the real return to capital is "too low". What is additionally quite striking is the fact that agents in the market economy may accumulate *more* human capital than what the different planners (the difference being based on how they discount the utility of the future generations) would have liked.

When will the market economy accumulate less human (and physical) capital than the planner? I show that answers to these sorts of questions depend on the specific nature of the consumption externality. For an example, if the consumption benchmark in society is largely influenced by the working population (the middle-aged), then underaccumulation is likely. Heuristically, this is a setting in which the Joneses are young and in trying to keep up with them, everyone wishes to consume more. The middle-aged achieve this by cutting down on saving. In a perfect-foresight equilibrium, this is consistent with a high return to capital and a low wage rate. Since the young borrow (at this high interest rate and in an environment of anticipated low wage rate) to fund their education, they per force borrow less and accumulate less human capital. A justification for public involvement in education is thereby created. While the informal argument presented above is useful for insight building, it deserves mention here that the exact discount rate used by the planner is also of critical importance – if, for example, the weight on future generations is high, the planner would set aside a lot of capital for future production.

From a modeling perspective, the paper closest to mine is Docquier et. al. (2007). The critical difference is that they assume the presence of a intergenerational human capital externality (which generates endogenous growth) but no consumption externality is present. The papers also differ in the way the two sets of solutions (one for competitive equilibrium and the other for the planner's solutions) are compared and the reference level of the planner's weight is defined (see section 3). Furthermore, they compare the levels of physical and human capital in a laissez-faire economy to that in the planner's solution. They find that in the presence of the human capital externality, the social weight assigned by the planner to future generations decides whether there is overaccumulation or underaccumulation of capital (both human and physical). In particular, when that weight is low (high) enough, the market always overaccumulates (underaccumulates). This characterization is no longer true in the present analysis. The exact nature of the consumption externality becomes crucial in determining the possibility of over and underaccumulation. There is an obvious implication of the Docquier et. al. (2007) result: if the planning weight on future generations is relatively high, the market underaccumulates and by implication, the case for public pensions is weakened. My results suggest that the case for public pensions may still be strong if consumption externalities of a certain kind are present.

The rest of the paper is organized as follows. Under section 2, I present the basic framework of the economy in subsection 2.1 and subsequently in subsection 2.2, I present the competitive set up while in subsection 2.3, a competitive equilibrium has been described. Planner's first best solutions in the presence of consumption externality have been characterized in subsection 2.4. In section 3, I present the main results of this paper. I work out the results using particular functional forms of a representative economy and present the numerical result in section 4. While section 5 concludes, the Appendix A contains the proofs of all the lemmas and propositions.

2 The Model

2.1 Primitives

I consider an economy consisting of an infinite sequence of three-period lived overlapping generations, an initial old generation, and an infinitely-lived government. In the first period of life (i.e., when they are young), agents borrow to invest in their education. They work and pay off their loans in the second period, and retire in the third period. Let $t = 1, 2, ...$ index time. The generation that works during period t is indexed by t. The population size of generation t is denoted by N_t. At any date t, the population consists of old, middle-aged, and young agents – the population size in period t is given by $N_{t-1} + N_t + N_{t+1}$. The population is assumed to grow at a gross rate of N or a net rate n i.e., $N = 1 + n$. The initial middle-aged agents are endowed with $K_1 > 0$ units of physical capital.

There is a single final good and it can either be consumed in the period it is produced, or it can be stored to yield capital the following period. Let K_t and H_t denote the aggregate levels of physical and human capital at date t respectively, and $h_t \equiv H_t/N_t$. The aggregate production function is given by $Y_t = F(K_t, H_t)$ where F is assumed to be homogeneous of degree 1. This assumption allows us to write $Y_t = H_t f(\overline{k_t})$, where, $\overline{k_t} \equiv \frac{k_t}{h_t}$ and $k_t = \frac{K_t}{N_t}$. The function f is assumed to be positive, strictly increasing and strictly concave in its argument, i.e., $f > 0$, $f' > 0$ and $f'' < 0$. For reasons of analytical tractability, capital is assumed to depreciate 100% between periods.

An agent belonging to generation t borrows an amount e_{t-1} from a perfect capital market and invests in education in period $t - 1$. The education they acquire in $t - 1$

translates into human capital in period t as described by

$$h_t = \phi(e_{t-1}), \tag{1}$$

where $\phi(\cdot)$ is a strictly increasing and a strictly concave function that satisfies the Inada conditions, i.e., $\phi'(\cdot) > 0$, $\phi''(\cdot) < 0$ with $\phi'(0) = \infty$, $\phi'(\infty) = 0$. As already discussed in the introduction, the construction of h_t in this paper ensures that the production of human capital is free from the effect of any kind of externality.

I now shed some light on the structure of human capital formation technology. In endogenous growth models like those used by Boldrin and Montes (2005), Docquier et al (2007) and others, the source of endogenous growth lies in the formulation of the human capital production function. In each period, human capital is generated using the human capital accumulated by their previous generation along with the amount the agent borrows from a perfect capital market. This specification ensures that there is an externality in the production of human capital since agents are born with some amount of human capital. Zhang (2003) has constructed a growth model where the human capital production function also includes the average human capital present in the economy. In this paper, I drop both the assumptions that per capita human capital production is affected by the average level of human capital and also that agents are born with some human capital. Thus I construct a human capital production function which is free from any type externality. Galor and Moav (2006) have also worked with a similar production function, albeit in a different context, where human capital production depends only on the amount of government subsidy available. I provide a perfect capital market to guarantee that any deviation from optimal human capital production is not due to the inexistence of a perfect source of borrowings.

Preferences of agents play an important role in the ensuing analysis. For simplicity, I assume the young at any date do not consume. Let c_t denote the middle-age consumption of a generation-t individual and let d_{t+1} denote her old-age consumption. I assume that an agent's utility depends upon her level of consumption when compared to a reference standard – her effective level of consumption. More precisely, lifetime utility of a generation-t agent is given by

$$u_t = u\left(\widehat{c}_t, \widehat{d}_{t+1}\right)$$

where $\widehat{c}_t$ and $\widehat{d}_{t+1}$ denote her effective level of consumption when middle aged and old respectively. This utility function is strictly increasing, strictly concave, and satisfies Inada conditions for both of its arguments. Following Alonso - Carrera et al. (2008) the effective levels of consumption of a generation-t agent is given by

$$\widehat{c}_t = c_t - \gamma v_t^m,$$

and

$$\widehat{d}_{t+1} = d_{t+1} - \delta v_{t+1}^o,$$

where $\gamma \in [0, 1)$ and $\delta \in [0, 1)$ measure the intensity of the consumption references, v_t^m and v_{t+1}^o, respectively. The consumption benchmarks in any period are assumed to be a weighted arithmetic average of the per-capita consumption of the two generations consuming in that period. Specifically,

$$v_t^m \equiv \frac{N_t c_t + \theta^m N_{t-1} d_t}{N_t + \theta^m N_{t-1}} = \left(\frac{N}{N + \theta^m}\right) c_t + \left(\frac{\theta^m}{N + \theta^m}\right) d_t,$$

where $\theta^m \in [0, 1]$ is the weight of a representative old agent's consumption in the specification of the middle-aged agent's consumption benchmark. Similarly

$$v_{t+1}^o \equiv \frac{\theta^o N_{t+1} c_{t+1} + N_t d_{t+1}}{\theta^o N_{t+1} + N_t} = \left(\frac{\theta^o N}{\theta^o N + 1}\right) c_{t+1} + \left(\frac{1}{\theta^o N + 1}\right) d_{t+1},$$

where $\theta^o \in [0,1]$ is the weight of a representative middle-aged agent's consumption in the specification of an old agent's consumption reference. Denoting $\frac{N}{N+\theta^m} \equiv \varepsilon^m$ and $\frac{\theta^o N}{\theta^o N+1} \equiv \varepsilon^o$ gives $v_t^m = \varepsilon^m c_t + (1-\varepsilon^m) d_t$ and $v_{t+1}^o \equiv \varepsilon^o c_{t+1} + (1-\varepsilon^o) d_{t+1}$.

The above formulation of preferences allow for both intragenerational consumption externality and also intergenerational consumption spillovers that is the possibility of a generation other than one's own to influence one's consumption decision. The strength of these influences depend on the deeper parameters θ^m and θ^o. Versions of these preferences have been used by Abel (2005).[4] It is worthwhile to note here that while ε^m $((1-\varepsilon^o))$ represents the degree to which a middle aged (old) agent keeps up with the others in her own generation, $(1-\varepsilon^m)$ and ε^o represents the degree to which the agent keeps up with the other generation. Barnett and Bhattacharya (2008) associate $\varepsilon^o > 0$ with 'rejuvenile' behavior, the old trying to keep up with their children. We now introduce some notions of externality based on the values of ε^m and ε^o. When both ε^m and ε^o are high, the weight that is given to the level of consumption of the middle aged (the working class) in the construction of the consumption reference is high, I call it a middle aged driven externality $[EM]$. Similarly, when both ε^m and ε^o are low, it is called an old driven externality $[EO]$. The case where ε^m is high but ε^o is low represents an economy where an agent's consumption externality is driven by her contemporary generation $[ECG]$. That is, the consumption externality of a middle aged agent is driven by the middle aged population whereas the the consumption externality of an old agent is driven by the old population. The condition when ε^o is high but ε^m is low represents an economy where an agent's consumption externality is driven by the other consuming generation $[EOG]$, i.e., the consumption externality of a middle aged agent is driven by the old population

[4] The form of externality used in this paper differs from the one used in Abel (2005). In Abel (2005) the reference consumption is of a multiplicative form and depends on the levels of consumption of two different generations at a particular time point. Furthermore the reference consumption level affects the utility of an agent's consumption multiplicatively. Here the reference consumption level has been constructed additively and unlike Abel (2005) effective consumption is determined by subtracting the reference consumption level from the actuals.

whereas the the consumption externality of an old agent is driven by the middle aged population. More specifically, I assume that 'high' means when the coefficients exceed the number $\frac{1}{2}$. Now I start with a competitive set up and then define the competitive equilibrium in this economy.

2.2 Trade

Middle-aged agents supply their human capital inelastically in competitive labor markets, earning a wage rate, w_t, at time t, where

$$w_t \equiv w(\bar{k}_t) = f(\bar{k}_t) - \bar{k}_t f'(\bar{k}_t) \tag{2}$$

and $w'(\bar{k}_t) > 0$. In addition, capital is traded in competitive capital markets, and earns a gross real return of R_{t+1} between t and $t+1$, where

$$R_{t+1} \equiv R\left(\bar{k}_{t+1}\right) = f'(\bar{k}_{t+1}) \tag{3}$$

with $R'\left(\bar{k}_{t+1}\right) < 0\ \forall t$ by the property of function f.

Parents in this economy are selfish and do not care for the education of their children. A generation-t agent borrows amount e_{t-1} in period $t-1$ at the gross interest rate R_t (also see Boldrin and Montes (2005), Docquier et al (2007)). Agent pays off this loan with income earned during period t. The income W_t of a middle-aged agent in period t is given by $W_t \equiv w_t h_t - R_t e_{t-1}$. A generation-$t$ agent's optimization problem can be written as:

$$\max_{s_t, e_{t-1}} u\left(\widehat{c}_t, \widehat{d}_{t+1}\right)$$

subject to

$$c_t = W_t - s_t,$$

$$d_{t+1} = R_{t+1} s_t,$$

and

$$c_t \geq \gamma v_t^m, \quad d_{t+1} \geq \delta v_{t+1}^o, \quad s_t \geq 0$$

where s_t denotes saving. Note that in a competitive set up, agents take v_t^m and v_{t+1}^o as parametrically given.

Assuming interior solutions and using (1), the solution to a generation-t agent's problem is characterized by the following optimality conditions

$$s_t : u_{\widehat{c}_t} = u_{\widehat{d}_{t+1}} R_{t+1} \Rightarrow \frac{u_{\widehat{c}_t}}{u_{\widehat{d}_{t+1}}} = R_{t+1} \tag{4}$$

and

$$e_{t-1} : w_t \, \phi'(e_{t-1}) = R_t \Rightarrow \phi'(e_{t-1}) = \frac{R_t}{w_t}. \tag{5}$$

Note that $s_t > 0 \Rightarrow W_t > 0$. The first optimality condition is straight forward which simply describes the optimum intertemporal consumption-saving decision of the agent. The second condition which represents the optimum expenditure towards education can also be looked at from another angle. The equation also guarantees that indeed the optimality is reached where there is no incentive of agents by increasing her education marginally, that is, where $\frac{\partial W_t}{\partial e_{t-1}} = w_t \, \phi'(e_{t-1}) - R_t = 0$ occurs. It can be seen that this equation is not directly affected by any externality in the consumption issue.

2.3 Competitive Equilibrium (CE)

Before I formally define the competitive equilibrium, I introduce the market clearing condition for this economy. While a part of the collective savings of the middle aged is used to finance education for the young, the remaining amount becomes capital stock for the next period. Hence the market clearing condition can be written as $N_t s_t =$

$K_{t+1} + N_{t+1}e_t$, which means

$$s_t = \overline{k}_{t+1}(1+n)h_{t+1} + (1+n)e_t. \tag{6}$$

I formally define competitive equilibrium below.

Definition 1 *Given k_1 and h_1, a competitive equilibrium for this economy is a sequence of consumption allocations $\{c_t, d_t\}_{t=1}^{\infty}$, allocations of saving, capital, and education expenses, $\{s_t; k_t; e_{t-1}\}_{t=1}^{\infty}$, and factor prices $\{w_t; R_t\}_{t=1}^{\infty}$ that solve the agents' optimization problem at each date, satisfy the market-clearing conditions, and the factor prices satisfy (2)-(3).*

Henceforth, I denote with superscript CE the competitive equilibrium outcomes. Using (2)-(3), I can rewrite (4)-(5) as

$$\frac{u^{CE}_{\widehat{c}_t}}{u^{CE}_{\widehat{d}_{t+1}}} = f'\left(\overline{k}^{CE}_{t+1}\right), \tag{7}$$

and

$$\phi'(e^{CE}_t) = \frac{f'\left(\overline{k}^{CE}_{t+1}\right)}{f\left(\overline{k}^{CE}_{t+1}\right) - \overline{k}^{CE}_{t+1}\, f'\left(\overline{k}^{CE}_{t+1}\right)}, \tag{8}$$

which, in turn, implies

$$e^{CE}_{t-1} = (\phi')^{-1}\left(\frac{R(\overline{k}^{CE}_t)}{w(\overline{k}^{CE}_t)}\right). \tag{9}$$

The income W_t can be re-written as

$$\begin{aligned} W_t &= w_t h_t - R_t e_{t-1} = w_t\phi(e_{t-1}) - w_t\phi'(e_{t-1})e_{t-1} \qquad (10) \\ &= w_t\phi(e_{t-1}(\overline{k}_t))[1-\eta] = w_t(\overline{k}_t)\phi(e_{t-1}(\overline{k}_t))[1-\eta(\overline{k}_t)], \end{aligned}$$

where η is the elasticity of ϕ with respect to e. Per capita saving at t by the working middle-aged is given by

$$s_t = W_t - c_t = s_t(W_t, d_t) = s_t(w_t(\overline{k}_t)\phi(e_{t-1}(\overline{k}_t))[1-\eta(\overline{k}_t)], R_t(\overline{k}_t)s_{t-1}). \tag{11}$$

In order to make the calculations needed to find (12) manageable, I assume that $\delta = 0$, i.e., the consumption of an old agent is not affected by any kind of externality.[5] Using equations (11) and (6) I get the equilibrium law of motion for the physical-to-human-capital ratio for the economy:

$$\begin{aligned} &s_t(w_t(\overline{k_t})\phi(e_{t-1}(\overline{k}_t))[1-\eta(\overline{k}_t)], R_t(\overline{k}_t)[\overline{k}_t(1+n)\phi(e_{t-1}(\overline{k}_t)) + (1+n)e_{t-1}(\overline{k}_t)]) \\ &= \overline{k}_{t+1}(1+n)\phi(e_t(\overline{k}_{t+1})) + (1+n)e_t(\overline{k}_{t+1}) \end{aligned} \tag{12}$$

All competitive equilibrium sequences $\{k_t\}$ and $\{h_t\}$ must satisfy (12).

A steady state equilibrium is a time-invariant sequence of c_t, d_t, s_t and e_t. In particular, in a steady state, a time invariant $\overline{k}^{CE}$ satisfies (12).

2.4 Social Planner's (SP) Solution

I now proceed to solve the planner's problem. The planner takes into account the cross-generational consumption externalities ignored by individual agents. It is easy to verify that the resource constraint for the planner can be written as:

$$h_t f(\overline{k}_t) = c_t + \frac{d_t}{1+n} + (1+n)(e_t + k_{t+1}). \tag{13}$$

The planner maximizes the sum of lifetime utilities of all the generations over the infinite horizon subject to the above resource constraint. I assume that generational

[5] However, no such restriction has been imposed on δ while presenting the main result. This assumption has been made only when finding the law of motion of physical-to-human capital ratio.

utility is discounted by a factor $\lambda \in (0,1)$. I refer to λ as the social weight that the planner attaches to the future generations. A higher λ implies that the social discount rate at which the planner devaluates the future generations is low. In the ensuing analysis, I will interpret different planners as each being indexed by their own λ. The Lagrangian for a generic planner's problem is as follows:

$$£ \equiv \sum_{t=0}^{\infty} \lambda^t \left\{ u(\widehat{c}_t, \widehat{d}_{t+1}) + q_t[h_t f(\overline{k}_t) - c_t - \frac{d_t}{1+n} - (1+n)(e_t + k_{t+1})] + p_t[\phi(e_t) - h_{t+1}] \right\}$$

where $\lambda^t q_t$ and $\lambda^t p_t$ are the multipliers associated with the resource constraint of the economy and human capital formation technology respectively, at date t.

The first order conditions with respect to c_t^{SP}, d_t^{SP}, e_t^{SP}, k_{t+1}^{SP} and h_{t+1}^{SP} are given below where the superscript SP denotes the planner's outcome:

$$c_t^{SP} : \lambda^t u_{\widehat{c}_t}^{SP}(1 - \gamma\varepsilon^m) - \lambda^{t-1} u_{\widehat{d}_t}^{SP} \delta \varepsilon^o - \lambda^t q_t = 0 \tag{14}$$

$$d_t^{SP} : -\lambda^t u_{\widehat{c}_t}^{SP}(1 - \varepsilon^m)\gamma + \lambda^{t-1} u_{\widehat{d}_t}^{SP}(1 - \delta(1 - \varepsilon^o)) - \lambda^t \frac{q_t}{1+n} = 0 \tag{15}$$

$$e_t^{SP} : -\lambda^t (1+n) q_t + \lambda^t p_t \phi'(e_t^{SP}) = 0 \tag{16}$$

$$\begin{aligned} k_{t+1}^{SP} \quad : \quad & \lambda^{t+1} q_{t+1} h_{t+1}^{SP} f'(\overline{k}_{t+1}^{SP}) \frac{1}{h_{t+1}^{SP}} - \lambda^t q_t (1+n) = 0 \\ \Rightarrow \quad & q_t(1+n) = \lambda q_{t+1} f'(\overline{k}_{t+1}^{SP}) \end{aligned} \tag{17}$$

$$h_{t+1}^{SP} \quad : \quad \lambda^{t+1} q_{t+1}[h_{t+1}^{SP} f'(\overline{k}_{t+1}^{SP})(-\frac{k_{t+1}^{SP}}{(h_{t+1}^{SP})^2}) + f(\overline{k}_{t+1}^{SP})] - \lambda^t p_t = 0 \tag{18}$$
$$\Rightarrow \quad \lambda^{t+1} q_{t+1}[f(\overline{k}_{t+1}^{SP}) - f'(\overline{k}_{t+1}^{SP})\overline{k}_{t+1}^{SP}] - \lambda^t p_t = 0.$$

From equations (16), (17) and (18), I have

$$\phi'(e_t^{SP}) = \frac{f'(\overline{k}_{t+1}^{SP})}{f(\overline{k}_{t+1}^{SP}) - f'(\overline{k}_{t+1}^{SP})\overline{k}_{t+1}^{SP}}. \tag{19}$$

Define the 'externality factor'

$$\Delta \equiv \frac{\delta\varepsilon^o + (1 - \delta(1 - \varepsilon^o))(1 + n)}{(1 - \gamma\varepsilon^m) + (1 - \varepsilon^m)\gamma(1 + n)}. \tag{20}$$

Then, from equations (14) and (15) I have

$$\frac{u_{\widehat{c}_t}^{SP}}{u_{\widehat{d}_t}^{SP}} = \frac{1}{\lambda}\Delta, \tag{21}$$

and using equations (14), (15) and (17), I can derive the following intertemporal relationship[6]:

$$\frac{u_{\widehat{c}_t}^{SP}}{u_{\widehat{d}_{t+1}}^{SP}} = f'(\overline{k}_{t+1}^{SP})\frac{\Delta}{1 + n}. \tag{22}$$

Combining above two equations, I get

$$\frac{u_{\widehat{d}_{t+1}}^{SP}}{u_{\widehat{d}_t}^{SP}} = \frac{1}{f'(\overline{k}_{t+1}^{SP})\lambda}(1 + n). \tag{23}$$

Note that at a steady state,

$$f'(\overline{k}^{SP}) = \frac{1}{\lambda}(1 + n) \tag{24}$$

[6]Algebraic calculations are shown in the supplementary materials.

holds which defines modified golden rule in this model. It is easy to verify that λ and $\overline{k}^{SP}$ are positively related, which in turn implies that if a planner gives more weight to future generations, there will be an increase in the steady state requirement of golden rule $\overline{k}^{SP}$.

3 Comparisons

In this section I compare a steady state of the competitive equilibrium with the planner's solution in terms of both physical and human capital. In order to be able to compare, I establish a common point around which this discussion will be meaningful. However, before I can do that, I will need to establish a few more results. All the proofs are given in the appendix.

First, it can be easily verified that $\dfrac{\partial \overline{k}_t^i}{\partial e_{t-1}^i} > 0$, $i = CE, SP$. Naturally the result also holds in a steady state. If I denote the elasticity of k_t with respect to e_{t-1} as ξ, then it is easy to check that for every k, h, e and t, $\xi > \eta$ holds, i.e., whenever the investment in education changes, the proportionate change in per capita physical capital dominates the proportionate change in human capital. This result is obtained by finding out the effect of social weight λ on the steady state values of $\overline{k}^{SP}$ and e^{SP}. It can be verified that both $\overline{k}^{SP}$ and e^{SP} increase as the social weight increases. Thus a planner who attaches a higher social weight to future generations guarantees a higher expenditure towards education. Since $\xi > \eta$ holds, an increase in social weight not only increases optimal per capita physical and human capital production but also the ratio of per capita physical to human capital. Thus, I have

Lemma 2 *At a steady state* $\frac{d\overline{k}^{SP}}{d\lambda} > 0$ *and* $\frac{de^{SP}}{d\lambda} > 0$ *hold for any* λ. *Also,* $\xi > \eta$ *holds for both* CE *and* SP.

Lemma 2 suggests that if the rate at which the planner discounts the future generation increases, expenditure towards education and hence the effective capital stock in the planner's solution will fall at a steady state. For an equivalent result in a growth set up, see Docquier et al (2007).

In order to establish a common ground where the comparison between the laissez-faire and planning solutions is meaningful, I assume there exists a λ (call it $\overline{\lambda}$) such that the optimality conditions generated by the planner's economy associated with this social weight is identical to that in the competitive economy. I also impose the additional restriction that $\overline{k}^{SP}(\overline{\lambda}) \in (0, (f')^{-1}(1+n))$, where $\overline{k}^{SP}(\overline{\lambda})$ represents $\overline{k}^{SP}$ associated with the weight $\overline{\lambda}$. Note that the values that $\overline{k}^{SP}$ can take, has an upper bound which is determined by the restriction on λ. In particular, as $\lambda \in (0,1)$, from equation (24), I have $\overline{k}^{SP} \longrightarrow 0$ as $\lambda \longrightarrow 0$ and on the other hand, $\overline{k}^{SP} \longrightarrow (f')^{-1}(1+n)$ when $\lambda \longrightarrow 1$. Since $\frac{d\overline{k}^{SP}}{d\lambda} > 0$, at the steady state $\overline{k}^{SP} \in (0, \overline{k}^{SP}_{\max})$ where $\overline{k}^{SP}_{\max} = (f')^{-1}(1+n)$.

At this point I should discuss the difference in the technique used to compare CE and SP by me with that in Docquier et al.(2007). Apart from the fact that their set up is different from ours, in Docquier et al.(2007), the benchmark social weights around which they have compared the laissez-faire and planning solutions are where the crucial variables (and hence capital stocks) are identical. That means, in their model there exist social weights for which the capital stocks in a planner's solution equal to that in a competitive economy. In contrast, in my model, there exists, under some restrictions, a particular social weight $\overline{\lambda}$, for which the planner's optimality conditions are exactly identical to that in the CE economy. However, in the absence of any consumption externality in the economy, planners optimality conditions as well as the allocations coincide with the laissez-faire one. I term this social weight $\overline{\lambda}$ a 'laissez-faire supported' social weight. The existence of such $\overline{\lambda}$ requires a restriction that $\overline{k}^{SP}(\overline{\lambda}) \in (0, \overline{k}^{SP}_{\max})$. I show if such a $\lambda = \overline{\lambda}$ exists then it must be unique. Hence the following lemma.

Lemma 3 *If there exists a laissez-faire supported social weight $\overline{\lambda}$, then it must be unique.*

There is another uniqueness result between $\overline{k}^{SP}$ and e^{SP} which is stated as Lemma 4 below. The lemma follows directly from equations (8) and (19). The result is quite expected as there is no externality present in the human capital formation function.

Lemma 4 *If there exists any* $\lambda = \widehat{\lambda}$ *so that* $\overline{k}^{SP}(\widehat{\lambda}) = \overline{k}^{CE}$ *holds, then the same* $\widehat{\lambda}$ *equates* $e^{SP}(\widehat{\lambda}) = e^{CE}$ *and vice-versa.*

Now I am in a position to state the main result of the paper.

Proposition 5 *At a laissez-faire supported social weight, competitive equilibrium underproduces both human and physical capital compared to the planning allocation if the externality factor is greater than unity, that is, if* $\Delta > 1$ *where* Δ *is defined in (20). On the other hand, competitive equilibrium overproduces both types of capital whenever the externality factor is less than unity, that is, if* $\Delta < 1$.

Thus according to the above proposition, if the social weight is $\overline{\lambda}$, we have $k^{CE} < k^{SP}$ and $h^{CE} < h^{SP}$ if $\Delta > 1$. However, if $\Delta < 1$, then $k^{CE} > k^{SP}$ and $h^{CE} > h^{SP}$.

I proceed to provide some intuition for these results. For simplicity, I assume that $n = 0$. It is interesting to note that $\Delta = 1$ in two different situations, a) when all externalities are absent, b) when $\frac{\delta}{\delta+\gamma}\varepsilon^{o} + \frac{\gamma}{\delta+\gamma}\varepsilon^{m} = \frac{1}{2}$ or $\frac{\delta}{\delta+\gamma}(1-\varepsilon^{o}) + \frac{\gamma}{\delta+\gamma}(1-\varepsilon^{m}) = \frac{1}{2}$, $\delta, \gamma \neq 0$. This later configuration is termed as "balanced" externality.

When externalities are present in both the periods, i.e., when an agent's consumption is affected by externalities when he is both middle aged and old, $\Delta \gtreqless 1 \Rightarrow \frac{\delta}{\delta+\gamma}\varepsilon^{o} + \frac{\gamma}{\delta+\gamma}\varepsilon^{m} \gtreqless \frac{1}{2} \Leftrightarrow \frac{\delta}{\delta+\gamma}(1-\varepsilon^{o}) + \frac{\gamma}{\delta+\gamma}(1-\varepsilon^{m}) \lesseqgtr \frac{1}{2}$. If there is no consumption externality when the agent is middle aged that is $\gamma = 0$, $\Delta \gtreqless 1 \Leftrightarrow \varepsilon^{o} \gtreqless \frac{1}{2}$. On the other hand, if there is no consumption externality when an agent is old, that is $\delta = 0$, we have $\Delta \gtreqless 1 \Leftrightarrow \varepsilon^{m} \gtreqless \frac{1}{2}$. Thus $\Delta > 1$ or $\Delta < 1$ occurs in many different situations. Observe that if an economy is characterized by EM, i.e., the externality in consumption is driven by the middle aged, we will always have $\Delta > 1$. This means that in this situation, if the planner's social weight is $\overline{\lambda}$, there will be an underaccumulation of both types of capital in the competitive equilibrium relative to the planner's solution. On the other hand, when an economy is classified as EO, if the social weight assigned by the planner is $\overline{\lambda}$, competitive equilibrium overproduces both types of capital when compared to a planner's solution . However, ambiguity arises when the economy is characterized by either ECG or EOG since in both these situations, Δ can either be less than or greater than unity.

The obvious question at this stage is how to decentralize the planner's solution. Any policy prescription based on the above results crucially depends on the social weight that the planner assigns to the future generations. In this model I allow the possibility that a planner's social weight may differ from $\overline{\lambda}$. Because of this change in λ, various interesting situations may arise. For an example, if a planner's social weight λ is smaller than $\overline{\lambda}$, even in an economy characterized by EM, competitive equilibrium may overproduce both types of capital compared to a planner's solution. Similarly, an economy which is characterized as EO may not necessarily overproduce both physical and human capital in a laissez-faire environment. It can be shown that if a planner assigns significant weight to the future generations, an economy characterized by EO may produce less of both types of capital. Thus policy prescription for an economy crucially depends on the social weight that the planner assigns to future generations. A complete characterization of all possible situations has been presented in the next two propositions.

Proposition 6 *Suppose the planner's social weight to the future generation* λ *exceeds the laissez-faire supported social weight* $\overline{\lambda}$*. Then competitive equilibrium underproduces both types of capital compared to the planning allocation if the externality factor is greater than or equal to unity, that is, if* $\Delta \geq 1$*. However, if the externality factor is less than unity, that is, if* $\Delta < 1$*, it is not obvious whether there will be underaccumulation or overaccumulation of capital in the competitive equilibrium relative to a planner's solution.*

The above proposition says that, given $\lambda > \overline{\lambda}$ holds, if $\Delta \geq 1$, then we have $h^{SP}(\lambda) > h^{CE}$ and $k^{SP}(\lambda) > k^{CE}$. On the other hand if $\Delta < 1$, the relationship between [$h^{SP}(\lambda)$ and h^{CE}] and [$k^{SP}(\lambda)$ and k^{CE}] becomes ambiguous.

Proposition 7 *Suppose the planner's social weight to the future generation λ is less than the laissez-faire supported social weight $\overline{\lambda}$. Then competitive equilibrium overproduces both types of capital compared to the planning allocation whenever the externality factor is less than or equal to unity, that is, if $\Delta \leq 1$. However, if the externality factor is greater than unity, that is, if $\Delta > 1$, it is not obvious whether there will be underaccumulation or overaccumulation of capital in the competitive equilibrium relative to a planner's solution.*

Thus, given $\lambda < \overline{\lambda}$ holds, if $\Delta \leq 1$, then $h^{SP}(\lambda) < h^{CE}$ and $k^{SP}(\lambda) < k^{CE}$. However, if $\Delta > 1$, the relationship between [$h^{SP}(\lambda)$ and h^{CE}] and [$k^{SP}(\lambda)$ and k^{CE}] becomes ambiguous.

I interpret the above phenomenon in the following paragraphs. All the different types of externalities and their corresponding effects can well be explained by the above results. I have clearly explained the situations where $\Delta \gtreqless 1$ holds.

Notice that when under $\lambda = \overline{\lambda}$, if I set $\delta = 0$, then $\Delta > 1 \Rightarrow \varepsilon^m > \frac{1}{2}$. That is, externality is present only when the agent is middle aged. Let me explain why there is an under accumulation of both the capitals in this scenario at the laissez-faire supported social weight. In this case since middle aged consumption has a high benchmark reference level, at a steady state, because of this high consumption, saving falls. This reduction in saving lowers the production of physical capital as a result of which the ratio of physical to human capital also falls at a given level of human capital. This fall in the ratio of physical to human capital makes cost of capital high as well as wage rate low at a given level of human capital. Thus, a reduction in saving through its effect on the ratio of physical to human capital makes cost of capital high and the wage rate low at a given level of human capital. Furthermore, since this high cost of capital and low wage rate reduce the incentive for young generation to borrow, this not only reduces the level of human capital accumulation but also reduces the production of physical capital even further. Therefore the ratio of physical to human capital ratio falls further since $\xi > \eta$ holds (see Lemma 2). When $\gamma = 0$, $\Delta > 1 \Rightarrow \varepsilon^o > \frac{1}{2}$. In this case there is no benchmark consumption when agents are middle aged. But since an old agent's affinity towards middle aged consumption is high, at the steady state, this increases her middle aged consumption too. Therefore the saving falls and the retrospective effects can be observed.

In the case of EM where there is underaccumulation of capital (under $\lambda = \overline{\lambda}$), observe that at a steady state, over-consumption by the middle aged agents leads to a reduction in savings which in turn by the virtue of the market clearing condition guarantees less per capita physical capital as well as less human capital production in the competitive equilibrium. Also, a fall in physical capital increases the cost of capital and decreases the wage rate. This in turn makes borrowing for human capital production less attractive. But if the weight assigned by a planner to the future generation is very low, it inversely affects (see Lemma 2) both physical and human capital production and thus leads to a lower level of physical and human capital at the social optimum. Therefore, in this situation where a planner's social weight is very low, an EM economy may overproduce both types of capital in a laissez-faire set up. Similarly, in the case of EO, equilibrium over-consumption by the old generation (under $\lambda = \overline{\lambda}$) leads to an increase in saving which in turn results in a greater production of both physical and human capital. But if the social weight is high, optimal capital production in a planner's economy increases (see Lemma 2) which in turn raises the possibility that the laissez-faire economy underproduces compared to a planner's economy. In all other cases, whether there is an overproduction or underproduction of capital in a competitive equilibrium relative to a social optima depends on the degree and type of externality. Notice that though I had started the discussion by finding out a specific $\lambda = \overline{\lambda}$. It is now clear from the results that we are capable of comparing the solutions of CE and SP for the entire range of the possible values of the social weight. The detail result of all the three propositions stated above are summarized in the table below.

Social weight	Externality factor	Relationship between CE and SP
$\overline{\lambda}$	$(1, > 1, < 1)$	(LF, UA, OA)[7]
$\lambda > \overline{\lambda}$	$(1, > 1, < 1)$	(UA, UA, AM)
$\lambda < \overline{\lambda}$	$(1, > 1, < 1)$	(OA, AM, OA)

4 An example

In this section I present my claims through an example. Consider the following representation of human capital and final good production technology respectively.

$$h_t = \phi(e_{t-1}) = Ae_{t-1}^{\alpha}\ ,\ \alpha \in (0,1),\ A > 0 \tag{25}$$

$$f(\overline{k}_t) = B\overline{k}_t^{\beta}\ ,\ \beta \in (0,1),\ B > 0 \tag{26}$$

Lifetime utility is log linear as shown below

$$u_t(\widehat{c}_t, \widehat{d}_{t+1}) = \log(c_t - \gamma v_t^m) + \rho \log(d_{t+1} - \delta v_{t+1}^o),\ \rho \in (0,1) \tag{27}$$

[7] LF, UA, OA and AM represent laissez-faire, underaccumulation, overaccumulation and ambiguous respectively. The table is to be read as follows: as an example, with reference to row 3, if the social weight λ is greater than the laissez-faire supported social weight $\overline{\lambda}$, competitive economy underaccumulates both human and physical capital if the externality factor Δ is greater than or equal to one. However, if the externality factor is less than one, the result is ambiguous.

where the intertemporal discount factor is ρ. From equation (9) I have

$$\begin{aligned} \phi'(e_{t-1}) &= \frac{R_t}{w_t} = \frac{\beta B\bar{k}_t^{\beta-1}}{(1-\beta)B\bar{k}_t^{\beta}} = \frac{\beta}{(1-\beta)\bar{k}_t} \\ &\Rightarrow \; A\alpha e_{t-1}^{\alpha-1} = \frac{\beta}{(1-\beta)\bar{k}_t} \Rightarrow e_{t-1} = \left(\frac{\beta}{A\alpha(1-\beta)\bar{k}_t}\right)^{\frac{1}{\alpha-1}}. \end{aligned} \tag{28}$$

Therefore I get

$$h_t = \phi(e_{t-1}) = A\left(\frac{\beta}{A\alpha(1-\beta)\bar{k}_t}\right)^{\frac{\alpha}{\alpha-1}}. \tag{29}$$

It is easy to check that the elasticity of ϕ with respect to e_{t-1}

$$\eta \equiv \frac{\phi'(e_{t-1})e_{t-1}}{\phi(e_{t-1})} = \alpha. \tag{30}$$

From equation (10) I get

$$\begin{aligned} W_t &= w_t(\bar{k}_t)\phi(e_{t-1}(\bar{k}_t))[1-\eta(\bar{k}_t)] = (1-\beta)B\bar{k}_t^{\beta} A(\frac{\beta}{A\alpha(1-\beta)\bar{k}_t})^{\frac{\alpha}{\alpha-1}}(1-\alpha) \\ &= BA(\frac{\beta}{A\alpha(1-\beta)})^{\frac{\alpha}{\alpha-1}}(1-\alpha)(1-\beta)\bar{k}_t^{\beta-\frac{\alpha}{\alpha-1}} \\ &= \underbrace{(1-\alpha)(1-\beta)AB\left(\frac{1}{A\alpha}\frac{\beta}{1-\beta}\right)^{\frac{\alpha}{\alpha-1}}}_{X}\bar{k}_t^{\beta-\frac{\alpha}{\alpha-1}} = X\bar{k}_t^{(\beta+\frac{\alpha}{1-\alpha})} \end{aligned} \tag{31}$$

and from equation (4) I have

$$\frac{1}{(c_t-\gamma v_t^m)} = \rho\frac{1}{(d_{t+1}-\delta v_{t+1}^o)}R_{t+1} \tag{32}$$

along with

$$c_t = w_t h_t - R_t e_{t-1} - s_t$$

$$d_{t+1} = R_{t+1}s_t.$$

Therefore, the consumption when they are middle aged can be calculated by

$$\begin{aligned} c_t &= w_t h_t - R_t e_{t-1} - s_t = W_t - \frac{d_{t+1}}{R_{t+1}} && (33) \\ &= W_t - \frac{\rho R_{t+1}(c_t - \gamma v_t^m) + \delta v_{t+1}^o}{R_{t+1}}. && (34) \end{aligned}$$

To avoid the difficulty in calculation, I set $\delta = 0$ which implies that there is no externality when the agent is old. Thus

$$c_t = \frac{W_t + \gamma\rho\,(1-\varepsilon^m)\,d_t}{\rho(1-\gamma\varepsilon^m)+1}.$$

Hence

$$s_t = W_t - c_t = W_t - \frac{W_t + \gamma\rho\,(1-\varepsilon^m)\,d_t}{\rho(1-\gamma\varepsilon^m)+1}$$

that is

$$\begin{aligned} s_t &= \frac{\rho(1-\gamma\varepsilon^m)X\overline{k}_t^{(\beta+\frac{\alpha}{1-\alpha})}}{\rho(1-\gamma\varepsilon^m)+1} - \frac{\gamma\rho\,(1-\varepsilon^m)\,d_t}{\rho(1-\gamma\varepsilon^m)+1} \text{ which implies} \\ s_t &= \frac{\rho(1-\gamma\varepsilon^m)X\overline{k}_t^{(\beta+\frac{\alpha}{1-\alpha})}}{\rho(1-\gamma\varepsilon^m)+1} - \frac{\gamma\rho\,(1-\varepsilon^m)\,R_t s_{t-1}}{\rho(1-\gamma\varepsilon^m)+1} \end{aligned}$$

using budget constraint of the agent. Since market clearing condition holds at any t, I replace s_{t-1} with the market clearing condition (6). Thus, given $R_t = f(\overline{k}_t)$, I have

$$\begin{aligned} s_t = &\frac{\rho(1-\gamma\varepsilon^m)X\overline{k}_t^{(\beta+\frac{\alpha}{1-\alpha})}}{\rho(1-\gamma\varepsilon^m)+1} - \\ &\frac{\gamma\rho\,(1-\varepsilon^m)\,\beta B\overline{k}_t^{\beta-1}\{(\overline{k}_t(1+n)A(\frac{\beta}{A\alpha(1-\beta)\overline{k}_t})^{\frac{\alpha}{\alpha-1}}) + ((1+n)(\frac{\beta}{A\alpha(1-\beta)\overline{k}_t})^{\frac{1}{\alpha-1}})\}}{\rho(1-\gamma\varepsilon^m)+1} \quad (35) \end{aligned}$$

Now equating the above equation with the market clearing condition again, I have,

$$\begin{aligned}
&\frac{\rho(1-\gamma\varepsilon^m)X\overline{k}_t^{(\beta-\frac{\alpha}{\alpha-1})}}{\rho(1-\gamma\varepsilon^m)+1} \\
&-\frac{\gamma\rho\left(1-\varepsilon^m\right)\beta B\overline{k}_t^{\beta-1}\{(\overline{k}_t(1+n)A(\frac{\beta}{A\alpha(1-\beta)\overline{k}_t})^{\frac{\alpha}{\alpha-1}})+((1+n)(\frac{\beta}{A\alpha(1-\beta)\overline{k}_t})^{\frac{1}{\alpha-1}})\}}{\rho(1-\gamma\varepsilon^m)+1} \\
= \;& \overline{k}_{t+1}(1+n)A(\frac{\beta}{A\alpha(1-\beta)\overline{k}_{t+1}})^{\frac{\alpha}{\alpha-1}}+(1+n)(\frac{\beta}{A\alpha(1-\beta)\overline{k}_{t+1}})^{\frac{1}{\alpha-1}}). \qquad (36)
\end{aligned}$$

For simplicity I assume that $n=0$ and define

$$Z_1 \equiv \frac{\rho(1-\gamma\varepsilon^m)\left(\frac{1}{A\alpha}\frac{\beta}{1-\beta}\right)^{\frac{\alpha}{\alpha-1}}(1-\alpha)(1-\beta)BA}{\rho(1-\gamma\varepsilon^m)+1}$$

$$Z_2 \equiv \frac{\gamma\rho\left(1-\varepsilon^m\right)\beta B}{\rho(1-\gamma\varepsilon^m)+1}$$

and

$$Z_3 \equiv \frac{A^{\frac{1}{1-\alpha}}((\frac{\beta}{\alpha(1-\beta)})^{\frac{\alpha}{\alpha-1}}+(\frac{\beta}{\alpha(1-\beta)})^{\frac{1}{\alpha-1}}))}{\rho(1-\gamma\varepsilon^m)+1}.$$

Then the path of $\{\overline{k}_t\}_{t=0}^{\infty}$, in particular, the equation (36), can be written as

$$(Z_1-Z_2Z_3)\overline{k}_t^{(\beta+\frac{\alpha}{1-\alpha})}=Z_3\overline{k}_{t+1}^{\frac{1}{1-\alpha}} \Rightarrow \overline{k}_{t+1}=(Z_1-Z_2Z_3)^{1-\alpha}\overline{k}_t^{\alpha+\beta-\alpha\beta}.$$

Given the specification of the economy, since $\alpha+\beta-\alpha\beta<1$ always holds, the path of $\{\overline{k}_t\}_{t=0}^{\infty}$ is a concave function. Then if $(Z_1-Z_2Z_3)>0$ the model guarantees a positive savings which in turn implies that

$$\frac{(1-\gamma\varepsilon^m)}{\gamma\left(1-\varepsilon^m\right)}>\frac{\alpha+\beta-\alpha\beta}{(1-\alpha)\left(1-\beta\right)}. \qquad (37)$$

Note that both $\frac{d}{d\gamma}\left(\frac{1-\gamma\varepsilon^m}{\gamma\left(1-\varepsilon^m\right)}\right)<0$ and $\frac{d}{d\varepsilon^m}\left(\frac{1-\gamma\varepsilon^m}{\gamma\left(1-\varepsilon^m\right)}\right)>0$. Observe when $\gamma=0$ the above condition reduces to $\frac{\alpha+\beta-\alpha\beta}{(1-\alpha)\left(1-\beta\right)}<\infty$, which is always true since

$\alpha, \beta \in (0,1)$. Further when $\varepsilon^m = 0$, I have $\dfrac{(1-\gamma\varepsilon^m)}{\gamma(1-\varepsilon^m)} = \frac{1}{\gamma}$ and when $\varepsilon^m \to 1$, I have $\dfrac{(1-\gamma\varepsilon^m)}{\gamma(1-\varepsilon^m)} \to \infty$. However when $\gamma \to 1$ and $\varepsilon^m = 1$, the ratio is undefined, but at the same time that is a meaningless model as $c_t = 0$. Along the steady state, I get

$$\overline{k}^{CE} = (Z_1 - Z_2 Z_3)^{\dfrac{1-\alpha}{1-\alpha+\beta-\alpha\beta}}.$$

I can compute $\overline{\lambda}$ by equating $f'(\overline{k}^{CE}) = \frac{\Delta}{\overline{\lambda}}$. Consequently $\overline{k}^{SP}$ can be calculated by $\overline{k}^{SP} = (f')^{-1}(\frac{1}{\overline{\lambda}})$ where in this example, $\overline{k}^{SP} = \left(\overline{\lambda}\beta B\right)^{\frac{1}{1-\beta}}$.

I now present a numerical example which clarifies my claims. Consider a particular economy where $\alpha = 0.2$, $\beta = 0.33$, $A = 10$, $B = 10$, $\rho = 0.9$. I compute the steady state values of $\left\{\overline{k}^{CE},\ k^{CE},\ h^{CE}\right\}$ and $\left\{\overline{k}^{SP},\ k^{SP},\ h^{SP}\right\}$ that are shown in Table 1 below for different parametric values of γ and ε^m.

Cases	γ	ε^m	$\overline{\lambda}$	Δ	$\left\{\overline{k}^{CE},\ h^{CE},\ k^{CE}\right\}$	$\left\{\overline{k}^{SP},\ h^{SP},\ k^{SP}\right\}$
(1)	0.9	0.9	0.3554	3.5714	{0.1898, 2.3434, 0.4447}	{1.2687, 3.7682, 4.7807}
(2)	0.9	0.2	0.0764	0.6494	{0.2437, 2.4946, 0.6079}	{0.1279, 2.1234, 0.2716}
(3)	0.2	0.9	0.5716	1.1905	{1.9877, 4.2158, 8.3799}	{2.5785, 4.4992, 11.6013}
(4)	1	0.5	0.0482	1.000	{0.2251, 2.4456, 0.5505}	{0.2251, 2.4456, 0.5505}

Table 1

It has already been established that the choice of $\overline{\lambda}$ depends on the value of Δ. But given $\Delta > 1$, if one planner's social weight λ is such that $\lambda < \overline{\lambda}$, I show that in this case, the relationship between $\left\{k^{CE},\ h^{CE}\right\}$ and $\left\{k^{SP},\ h^{SP}\right\}$ is not unidirectional. A similar situation arises when a planner's social weight λ is such that $\lambda > \overline{\lambda}$ and $\Delta < 1$. I show this numerically in Table 2 and in two figures, Figure 1 and Figure 2, below. I pick case (1) from Table 1 where the value of Δ is greater than 1 and $\overline{\lambda} = 0.3554$. I choose two possible values of λ, both being less than $\overline{\lambda} = 0.3554$. This result is shown as case $(1')$ in Table 2. When $\lambda = 0.34$, competitive equilibrium underproduces and when $\lambda = .034$, competitive equilibrium overproduces both types of capital. Similarly, in $(2')$, where Δ is less than 1, I choose two possible values of λ both of which are greater than the corresponding $\overline{\lambda} = 0.0764$ (see case (2) in Table 1). Observe that in one situation ($\lambda = 0.09$), competitive equilibrium overproduces while in the other ($\lambda = 0.9$), it underproduces both types of capital compared to the planner's economy.

Cases	γ	ε^m	λ	Δ	$\left\{\overline{k}^{CE},\ h^{CE},\ k^{CE}\right\}$	$\left\{\overline{k}^{SP},\ h^{SP},\ k^{SP}\right\}$
$(1')$	0.9	0.9	0.34	3.5714	{0.1898, 2.3434, 0.4447}	{1.1875, 3.7063, 4.4011}
			0.034		{0.1898, 2.3434, 0.4447}	{0.0382, 1.5697, 0.0600}
$(2')$	0.9	0.2	0.09	0.6494	{0.2437, 2.4946, 0.6079}	{0.1633, 2.2571, 0.3687}
			0.90		{0.2437, 2.4946, 0.6079}	{5.0770, 5.3296, 27.0582}

Table 2

Insert [Figure1 : Case $(1')$] and [Figure2 : Case $(2')$] here.

5 Conclusion

Governments commonly intervene in education, typically in the form of education loans and subsidies. The standard rationale for such intervention is a human capital externality: for the same effort, people learn more if they are around smart people. The intergenerational counterpart of this observation is that a smart generation produces a smarter future generation. This paper make the case for government intervention in education even when no human capital externalities are present.[8] A neoclassical overlapping

[8] There are alternative ways to achieve efficiency. I do not claim that provision of education subsidies is the unique efficient way. For example, a tax or a subsidy (enough to correct the effects of the consumption externality) on capital income, along with a lump-sum transfer can implement the planner's

generations model of human and physical capital accumulation is studied. Children borrow from perfect capital markets to fund education expenses. When middle-aged, they earn income from human capital, and save in the form of physical capital. Agents are assumed to care not just about the level of their consumption but how that compares to those (from among their peers and the other consuming generation) living around them. Such an intra and intergenerational consumption externality is responsible for the possibility that agents may over or underaccumulate human and physical capital relative to a planner, thereby justifying public intervention.

solution; also see Richter and Braun (2010). While such alternatives somewhat weaken the case for direct government involvement in education, it can be argued that many countries find it considerably difficult to employ instruments such as a capital-income tax, and choose direct government involvement in education as a easier and simpler alternative.

Appendix

Proof of Lemma 2

The proof is simple and straightforward. Taking the total derivative of either equation (8) or (19) gives

$$\phi''(e_t)[f(\overline{k}_{t+1}) - f'(\overline{k}_{t+1})\overline{k}_{t+1}]de_t + \phi'(e_t)[-\overline{k}_{t+1}f''(\overline{k}_{t+1})]d\overline{k}_{t+1} - f''(\overline{k}_{t+1})d(\overline{k}_{t+1}) = 0$$

$$\Rightarrow \frac{d\overline{k}_{t+1}}{de_t} = \frac{\phi''(e_t)[f(\overline{k}_{t+1}) - f'(\overline{k}_{t+1})\overline{k}_{t+1}]}{f''(\overline{k}_{t+1})[1+\phi'(e_t)\overline{k}_{t+1}]}. \tag{A.1}$$

Given the property of f and ϕ, I clearly have $\frac{d\overline{k}_{t+1}}{de_t} > 0 \ \forall t$ for both the CE and SP. Notice that since $\frac{d\overline{k}_t}{de_{t-1}} = \frac{\frac{\partial k_t}{\partial e_{t-1}}h_t - \frac{\partial h_t}{\partial e_{t-1}}k_t}{{h_t}^2}$, $\frac{d\overline{k}_t}{de_{t-1}} > 0 \Rightarrow \frac{\partial k_t}{\partial e_{t-1}}h_t - \frac{\partial h_t}{\partial e_{t-1}}k_t > 0$ $\Rightarrow \frac{\partial k_t}{\partial e_{t-1}}\frac{e_{t-1}}{k_t} - \frac{\partial h_t}{\partial e_{t-1}}\frac{e_{t-1}}{h_t} > 0$ and thus $\xi > \eta$.

To prove $\frac{d\overline{k}^{SP}}{d\lambda} > 0$, at the steady state, I take the total derivative of equation (24) which implies that

$$f'(\overline{k}^{SP})d\lambda + \lambda f''(\overline{k}^{SP})d\overline{k}^{SP} = 0 \Rightarrow \frac{d\overline{k}^{SP}}{d\lambda} = \frac{-f'(\overline{k}^{SP})}{\lambda f''(\overline{k}^{SP})} > 0. \tag{A.2}$$

Using (A.1) and (A.2) I clearly have $\frac{de^{SP}}{d\lambda} > 0$.

Proof of Lemma 3

Suppose not. Let there exist $\overline{\lambda}$ and $\widetilde{\lambda}$, $\overline{\lambda} \neq \widetilde{\lambda}$ so that $\frac{u^{CE}_{\widehat{c}_t}}{u^{CE}_{\widehat{d}_{t+1}}} = \frac{u^{SP}_{\widehat{c}_t}}{u^{SP}_{\widehat{d}_{t+1}}}$ holds for both $\overline{\lambda}$ and $\widetilde{\lambda}$. Without loss of generality, assume that $\overline{\lambda} > \widetilde{\lambda}$. But if $\overline{\lambda} > \widetilde{\lambda}$, using Lemma 2, I have $\overline{k}^{SP}(\overline{\lambda}) > \overline{k}^{SP}(\widetilde{\lambda})$. But on the other hand, both $f'\left(\overline{k}^{CE}\right) = f'\left(\overline{k}^{SP}(\overline{\lambda})\right)\Delta$ and $f'\left(\overline{k}^{CE}\right) = f'\left(\overline{k}^{SP}(\widetilde{\lambda})\right)\Delta$ imply that $\overline{k}^{SP}(\overline{\lambda}) = \overline{k}^{SP}(\widetilde{\lambda})$ given $f' > 0$. Hence the contradiction.

Proof of Lemma 4

The proof is straightforward. It can directly be constituted from equation (19). Given $\overline{k}^{SP}(\widehat{\lambda}) = \overline{k}^{CE}$, I have

$$\begin{aligned} \phi'(e^{SP}(\widehat{\lambda})) &= \frac{f'(\overline{k}^{SP}(\widehat{\lambda}))}{f(\overline{k}^{SP}(\widehat{\lambda})) - f'(\overline{k}^{SP}(\widehat{\lambda}))(\overline{k}^{SP}(\widehat{\lambda}))} = \frac{f'(\overline{k}^{CE})}{f(\overline{k}^{CE}) - f'(\overline{k}^{CE})(\overline{k}^{CE})} = \phi'(e^{CE}) \\ &\Rightarrow \; e^{SP}(\widehat{\lambda}) = e^{CE} \text{ by the property of } \phi(*). \end{aligned}$$

Proof of Proposition 5

Note that when $\Delta > 1$, I have $f'\left(\overline{k}^{CE}\right) > f'\left(\overline{k}^{SP}(\overline{\lambda})\right)$ which implies that $\overline{k}^{CE} < \overline{k}^{SP}(\overline{\lambda})$. But $\overline{k}^{CE} < \overline{k}^{SP}(\overline{\lambda}) \Rightarrow e^{SP} > e^{CE} \Rightarrow h^{SP}(\overline{\lambda}) > h^{CE}$. Therefore $h^{SP}(\overline{\lambda}) > h^{CE}$ and $\overline{k}^{CE} < \overline{k}^{SP}(\overline{\lambda})$ gives $k^{SP}(\overline{\lambda}) > k^{CE}$. Similarly $k^{CE} > k^{SP}$ and $h^{CE} > h^{SP}$ if $\Delta < 1$. And definitely $k^{CE} = k^{SP}$ and $h^{CE} = h^{SP}$ hold if $\Delta = 1$.

Proof of Proposition 6

Given $\lambda > \overline{\lambda}$, I must have $\overline{k}^{SP}(\lambda) > \overline{k}^{SP}(\overline{\lambda})$ since I have already proved $\frac{d\overline{k}^{SP}}{d\lambda} > 0$ in lemma 2. But if $\Delta > 1$ holds, I must have $\overline{k}^{SP}(\lambda) > \overline{k}^{SP}(\overline{\lambda}) > \overline{k}^{CE}$. Since lemma 2 holds, this in turn implies that $h^{SP}(\lambda) > h^{CE}$ and thus I also have $k^{SP}(\lambda) > k^{CE}$. Also note that if $\Delta = 1$, $\overline{k}^{SP}(\lambda) > \overline{k}^{SP}(\overline{\lambda}) = \overline{k}^{CE}$. This by the virtue of lemma 2 means that $h^{SP}(\lambda) > h^{CE}$ and therefore $k^{SP}(\lambda) > k^{CE}$. But when $\Delta < 1$, $\overline{k}^{CE} > \overline{k}^{SP}(\overline{\lambda})$ holds along with $\overline{k}^{SP}(\overline{\lambda}) < \overline{k}^{SP}(\lambda)$ which proves the relationship is ambiguous. Hence the proof.

Proof of Proposition 7

Similar to the proof of Proposition 6.

References

[1] Abel, A., 2005. Optimal Taxation When Consumers Have Endogenous Benchmark Levels of Consumption. The Review of Economic Studies, 72, 1 pp 21-42.

[2] Acemoglu, D., Angrist, J., 2000. How Large Are Human-Capital Externalities? Evidence from Compulsory Schooling Laws, NBER Macroeconomics Annual, pp 9-59

[3] Alonso -Carrera, J., Caballe, J., Raurich, X., 2008. Estate taxes, consumption externalities, and altruism. Journal of Public Economics 92, pp 1751-64

[4] Barnett, R., Bhattacharya, J., 2008. Rejuveniles and Growth. European Economic Review 52, pp. 1055-1071

[5] Becker, G. S., 1975. Human capital: A theoretical and empirical analysis with special relevance to education, second edition (NBER, New York)

[6] Boldrin, M., Montes, A., 2005. The Intergenerational State Education and Pensions. Review of Economic Studies 72, pp 651-664.

[7] Card, D., 1999. The Causal Effect of Education on Earnings. Handbook of Labor Economics (Edited by Orley Ashenfelter and David Card) 3A, Amsterdam, North Holland

[8] Ciccone, A., Perri, G., 2006. Identifying Human-Capital Externalities: Theory with Applications, Review of Economic Studies 73, pp 381-412

[9] Dee, T. S. , Are there civic returns to education?, Journal of Public Economics, 88, pp 1697-1720

[10] Diamond, P. A., 1965. National debt in a neoclassical growth model, American Economic Review. 55, pp1126–1250

[11] Docquier, F., Paddison, O., Pestieau, P., 2007. Optimal accumulation in an endogenous growth setting with human capital. Journal of Economic Theory 134, pp 361-378

[12] Durlauf, S. N., Fafchamps, M., 2004. Social Capital. Handbook of Economic Growth, Steven Durlauf and Philippe Aghion (eds.)

[13] Galor, O., Moav, O., 2006. Das Human Kapital: A Theory of the Demise of the Class Structure, Review of Economics Studies, 73, 85-117

[14] Heckman, James J., 2000. Policies to Foster Human Capital. Research in Economics 51, pp 3-56

[15] Heckman, James J. and Peter J. Klenow,1998. Human Capital Policy. Policies to Promote Capital Formation (Edited by Michael Boskin), Stanford, CA, Hoover Institution

[16] Lange, F., Topel, R., The Social Value of Education. Handbook of the Economics of Education (Edited by E Hanushek and F. Welch), Amsterdam, Elsevier, 2006, pp 459-509

[17] Layard, R., 1980. Human satisfactions and public policy. The Economic Journal 90, pp 737-750

[18] Lucas, R. E., 1988. On the mechanics of economic development. Journal of Monetary Economics 21, pp 3-42

[19] Putnam, R. D., Leonardi, R., Nanetti, R. Y., 1993. Making Democracy Work: civic traditions in Modern Italy. Princeton: Princeton University Press.

[20] Richter F. W. , Braun C., Efficient subsidization of human capital accumulation with overlapping generations and endogenous growth, CESIFO working paper, 2010.

[21] Rudd, J., 2000. Empirical Evidence on Human Capital Spillovers, Finance and Economics Discussion Series 2000-46, Washington: Board of Governors of the Federal Reserve System

[22] Turner, C., Tamura, R., Mulholland, S.,E., Baier, S. , 2007. Education and Income of the States of the United States. Journal of Economic Growth 12, pp 1840-2000

[23] Yamarik, S. J., 2008. Estimating Returns to Schooling from State-Level Data: A Macro-Mincerian Approach. The B.E. Journal of Macroeconomics 8, 1 (Contributions), Article 23.

[24] Zhang, J., 2003. Optimal debt, endogenous fertility, and human capital externalities in a model with altruistic bequests. Journal of Public Economics 87, pp 1825-1835

Supplementary materials

Proof of $\left(\dfrac{u_{\widehat{c}_t}}{u_{\widehat{d}_{t+1}}}\right)^{SP} = f'(\overline{k}^{SP}_{t+1})\frac{\Delta}{1+n}$ [Equation (22) of the manuscript]**:**

From (14) of the manuscript I have

$$\lambda^t u^{SP}_{\widehat{c}_t}(1-\gamma\varepsilon^m) - \lambda^{t-1}u^{SP}_{\widehat{d}_t}\delta\varepsilon^o = \lambda^t q_t \tag{B.1}$$

and from (15) of the manuscript I get

$$-\lambda^t u^{SP}_{\widehat{c}_t}(1-\varepsilon^m)\gamma + \lambda^{t-1}u^{SP}_{\widehat{d}_t}(1-\delta(1-\varepsilon^o)) = \lambda^t \frac{q_t}{1+n}. \tag{B.2}$$

Using the above two equations (by replacing q_t of the first equation by the second expression of q_t) I have

$$\begin{aligned}
& \lambda^t u^{SP}_{\widehat{c}_t}(1-\gamma\varepsilon^m) - \lambda^{t-1}u^{SP}_{\widehat{d}_t}\delta\varepsilon^o \\
= \; & -\lambda^t u^{SP}_{\widehat{c}_t}(1-\varepsilon^m)\gamma(1+n) + \lambda^{t-1}u^{SP}_{\widehat{d}_t}(1-\delta(1-\varepsilon^o))(1+n) \\
\Rightarrow \; & \lambda^t u^{SP}_{\widehat{c}_t}\{(1-\gamma\varepsilon^m) + \lambda^t u^{SP}_{\widehat{c}_t}(1-\varepsilon^m)\gamma(1+n)\} \\
= \; & \lambda^{t-1}u^{SP}_{\widehat{d}_t}(1-\delta(1-\varepsilon^o))(1+n) + \lambda^{t-1}u^{SP}_{\widehat{d}_t}\delta\varepsilon^o \\
\Rightarrow \; & \lambda u^{SP}_{\widehat{c}_t}\{(1-\gamma\varepsilon^m) + (1-\varepsilon^m)\gamma(1+n)\} \\
= \; & u^{SP}_{\widehat{d}_t}\{(1-\delta(1-\varepsilon^o))(1+n) + \delta\varepsilon^o\} \\
\Rightarrow \; & \frac{u^{SP}_{\widehat{c}_t}}{u^{SP}_{\widehat{d}_t}} = \frac{(1-\delta(1-\varepsilon^o))(1+n) + \delta\varepsilon^o}{(1-\gamma\varepsilon^m) + (1-\varepsilon^m)\gamma(1+n)}.
\end{aligned} \tag{B.3}$$

Note that from (B.1), I can write

$$-\lambda u^{SP}_{\widehat{c}_t} = \frac{1}{(1-\gamma\varepsilon^m)}\left(-\lambda q_t - u^{SP}_{\widehat{d}_t}\delta\varepsilon^o\right). \tag{B.4}$$

I now substitute $-\lambda^t u^{SP}_{\widehat{c}_t}$ in the equation (B.2) by the above expression of $-\lambda u^{SP}_{\widehat{c}_t}$ (equation B.4) and I get the following

$$\left(-\lambda q_t - u^{SP}_{\widehat{d}_t}\delta\varepsilon^o\right)\frac{(1-\varepsilon^m)\gamma}{(1-\gamma\varepsilon^m)} + u^{SP}_{\widehat{d}_t}(1-\delta(1-\varepsilon^o)) = \lambda\frac{q_t}{1+n}$$

$$\Rightarrow -u^{SP}_{\widehat{d}_t}\frac{\delta\varepsilon^o(1-\varepsilon^m)\gamma}{(1-\gamma\varepsilon^m)} + u^{SP}_{\widehat{d}_t}(1-\delta(1-\varepsilon^o)) = \lambda q_t\left(\frac{1}{1+n} + \frac{(1-\varepsilon^m)\gamma}{(1-\gamma\varepsilon^m)}\right)$$

$$\Rightarrow u^{SP}_{\widehat{d}_t}\left\{1-\delta(1-\varepsilon^o) - \frac{\delta\varepsilon^o(1-\varepsilon^m)\gamma}{(1-\gamma\varepsilon^m)}\right\} = \lambda q_t\left(\frac{1}{1+n} + \frac{(1-\varepsilon^m)\gamma}{(1-\gamma\varepsilon^m)}\right)$$

$$\Rightarrow u^{SP}_{\widehat{d}_t} = \lambda q_t \underbrace{\frac{\left(\frac{1}{1+n} + \frac{(1-\varepsilon^m)\gamma}{(1-\gamma\varepsilon^m)}\right)}{\left\{1-\delta(1-\varepsilon^o) - \frac{\delta\varepsilon^o(1-\varepsilon^m)\gamma}{(1-\gamma\varepsilon^m)}\right\}}}_{\Pi}$$

$$\Rightarrow u^{SP}_{\widehat{d}_t} = \lambda q_t \Pi \text{ where } \Pi = \frac{\left(\frac{1}{1+n} + \frac{(1-\varepsilon^m)\gamma}{(1-\gamma\varepsilon^m)}\right)}{\left\{1-\delta(1-\varepsilon^o) - \frac{\delta\varepsilon^o(1-\varepsilon^m)\gamma}{(1-\gamma\varepsilon^m)}\right\}}. \tag{B.5}$$

But note that from (17) of the manuscript, I have

$$q_t = \frac{\lambda f'(\overline{k}^{SP}_{t+1})}{(1+n)} q_{t+1}. \tag{B.6}$$

Thus combining (B.5) and (B.6) I have

$$u^{SP}_{\widehat{d}_t} = \frac{\lambda^2 \Pi f'(\overline{k}^{SP}_{t+1})}{(1+n)} q_{t+1}. \tag{B.7}$$

Since (B.5)$\Rightarrow u^{SP}_{\widehat{d}_{t+1}} = \lambda q_{t+1}\Pi$ I have

$$u^{SP}_{\widehat{d}_{t+1}} = \lambda q_{t+1}\Pi = \lambda\frac{q_t(1+n)}{\lambda f'(\overline{k}^{SP}_{t+1})}\Pi \tag{B.8}$$

substituting q_{t+1} by the expression in (B.6). But in (B.5), I have an expression for q_t in terms of $u^{SP}_{\widehat{d}_t}$. Thus in the above expression (B.8) I replace the term q_t using equation (B.5). This implies

$$\begin{aligned} u^{SP}_{\widehat{d}_{t+1}} &= \lambda q_{t+1}\Pi = \lambda \frac{\frac{u^{SP}_{\widehat{d}_t}}{\lambda\Pi}(1+n)}{\lambda f'(\overline{k}^{SP}_{t+1})}\Pi \\ &= \frac{(1+n)}{f'(\overline{k}^{SP}_{t+1})} u^{SP}_{\widehat{d}_t} \\ &\Rightarrow \; u^{SP}_{\widehat{d}_t} = \frac{f'(\overline{k}^{SP}_{t+1})}{(1+n)} u^{SP}_{\widehat{d}_{t+1}}. \end{aligned} \tag{B.9}$$

Thus using (B.3) and (B.9) I have

$$\frac{u^{SP}_{\widehat{c}_t}}{u^{SP}_{\widehat{d}_{t+1}}} = \underbrace{\frac{(1-\delta(1-\varepsilon^o))(1+n)+\delta\varepsilon^o}{(1-\gamma\varepsilon^m)+(1-\varepsilon^m)\gamma(1+n)}}_{\Delta} \frac{f'(\overline{k}^{SP}_{t+1})}{(1+n)} = \frac{\Delta}{(1+n)} f'(\overline{k}^{SP}_{t+1}).$$

Hence the proof.

Figures

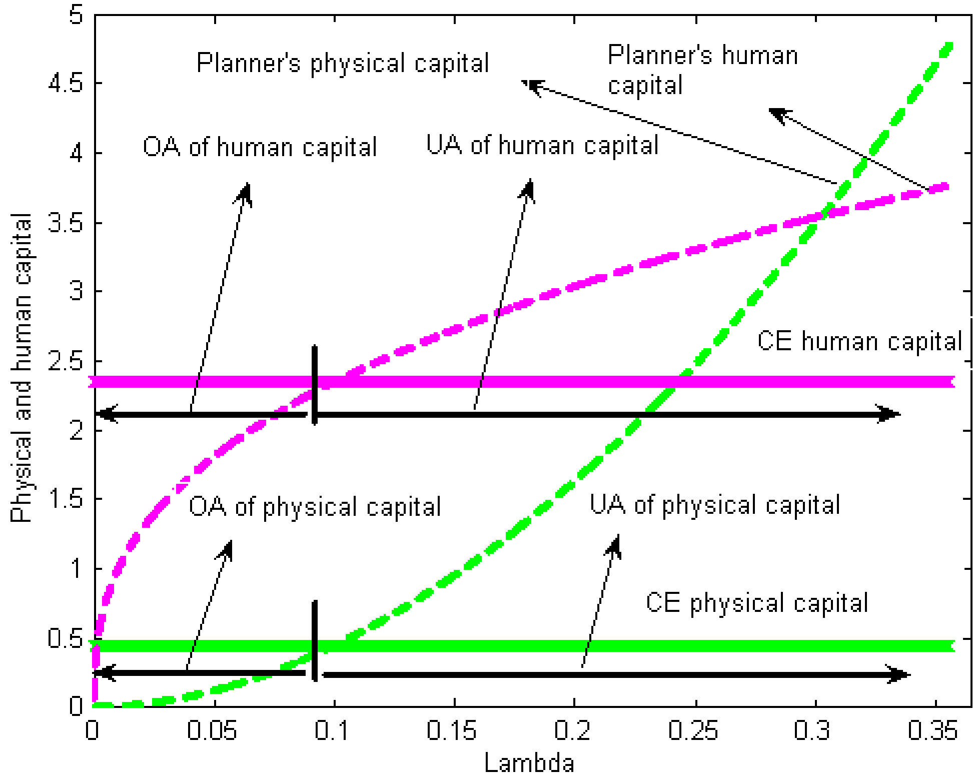

Figure 1: Case (1′)

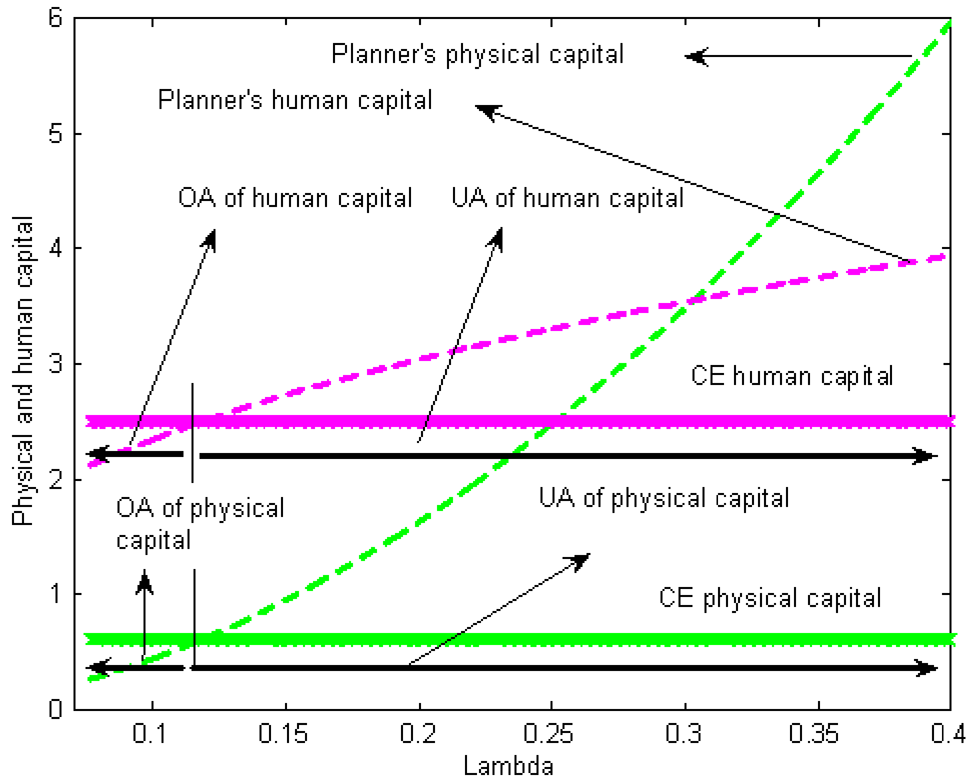

Figure 2: Case (2′)

CHAPTER 4. STOCHASTIC GROWTH, EDUCATION SUBSIDY AND INVARIANT DISTRIBUTION

Modified from a paper written jointly with Krishna B. Athreya

1 Introduction

The topic on economic growth is a widely researched and a well documented one. Economic growth models have been studied in both non-stochastic as well as in stochastic setups. Brock and Mirman have extended the standard one sector neoclassical optimal growth problem in a deterministic setup (see Ramsey (1928), Cass (1965), Koopmans (1965)) to a stochastic environment. The uncertainty in this economy is generated by the productivity shock. They show that the existence, uniqueness, and stability results of the deterministic case can also be realized in a stochastic model under similar assumptions. Significant works in this area have also been done by Mirman (1972, 1973), Mirman and Zilcha (1975) (also see Bhattacharya and Majumdar (2007)), Brock and Majumdar (1978), Majumdar and Zilcha (1987), Razin and Yahav (1979), Donaldson and Mehra (1983), Stokey et al. (1989), Hopenhayn and Prescott (1992), Amir (1997), Nishimura and Stachurski (2005) and Stachurski (2002, 2003) among others. The issues of existence, uniqueness and the stability of equilibrium in overlapping generations model have been studied by Galor and Ryder (1989). Stachurski (2003) extended this analysis in a stochastic environment. The proof of existence of steady state for a simple decentralized growth model with random shock has been studied by Laitner (1981) too.

In this paper, we present an endogenous growth model in a stochastic framework. As in De la Croix and Michel (2002), Boldrin and Montes (2004, 2005), Docquier et al (2007), Bishnu (2009) among others, agents in our economy live for three periods. In period one they acquire education, in period two they work and save for the future and in period three they consume out of their savings. In this model, the production of the final good not only uses physical capital but also human capital. The stochastic nature of the model arises from the fact that production is subject to a period-wise shock. Furthermore, human capital in each period is produced using the human capital of the previous period and also an education subsidy which is made available through an intergenerational tax arrangement. The income of the working class is subject to a proportional tax and the tax revenue is used to subsidize education. Given this benchmark set-up, the novelty of this paper lies in two aspects. First, we construct a growth model in a stochastic environment where more than one sector is present and contributes to the production of the final good. We provide a concept of long run equilibrium in this economy by defining an appropriate version of a balanced growth path for a stochastic environment. The concept of balanced growth path has been extensively studied in models with economic growth. Imposition of this economic condition is not only theoretically appealing but also empirically supported. In a non-stochastic setup, a balanced growth path is defined where the crucial variables grow at a constant rate in line with the Kaldor facts of economic growth. Kaldor had emphasized the phenomenon that the growth rate of output, factor-output ratio, real interest rate, share of labor and capital in income are roughly constant over time. This constancy of crucial ratios is an important characterization of an economy. However, in a stochastic setup, the stocks of both physical and human capital are random variables. So we define the global equilibrium of this economy in the long run as where the *ratio* of the two capital stocks converges to a steady state distribution. We name it the *Stochastic Balanced Growth Path*. We show that under some reasonable conditions, this steady state equilibrium exists and is unique.

Second, we study the tax rate (and hence subsidy) at which the long run equilibrium of our benchmark economy is able to achieve the first best. As mentioned in the previous paragraph, one alternative to fund education is an education subsidy. This arrangement involves the imposition of a proportional tax by the government on the working agents. It then uses the tax revenue to subsidize education. To compare this with the first best, we infuse an alternative source of borrowing funds for education, namely a perfect capital market. We investigate whether our benchmark economy has the capability to follow the same growth path as an economy with a perfect capital market. We have shown that the stochastic steady state exists and is unique for this latter economy too. We also show that for a particular tax rate, the stochastic balanced growth path of our benchmark economy converges to the same steady state distribution as the economy with a perfect capital market[1]. Using this result we further confirm that at this particular tax rate, the factor price ratios under the two different funding schemes also converge to the same steady state distribution. Moreover we also prove that for any tax rate lower than this particular tax rate, the stochastic balanced growth path first order stochastically dominates the balanced growth path of complete market allocation at the steady state. The assumption of boundedness of the shock is helpful to prove this result. We should mention at this point that in this paper we do not aim to achieve either a complete market allocation or economic efficiency from our benchmark model using a tax scheme[2]. Our main purpose is to show that in a stochastic framework, an unique steady exists under the intergenerational tax regime and this steady state distribution coincides with the distribution of the first best arrangement. This makes it possible to compare the two economies along an invariant steady state. A separate study may be initiated to handle

[1]Note that in a non-stochastic framework, this does not imply that the allocations in these two economies are same (Boldrin and Montes (2005)).

[2]In a non-stochastic setup, there are some studies which raise some concerns. In the framework of overlapping generation model, if the present generation is selfish, why should they invest in an asset which is valuable only to the next genartion (see Kotlikoff, Persson and Svenson (1988)). A social contract may be designed to handle the issue of public education along with public pension (see Boldrin (1992), Boldrin and Rustichini (2000)). Also, there is a possibility that our model may generate moral hazard problem; young may receive funds for education but may refuse to invest when they are middle aged and this can lead to a potential market faliure.

the issue of efficiency in our framework.

2 Background

We consider an economy consisting of an infinite sequence of three-period lived generations, an initial old generation, and an infinitely-lived government. In the first period (i.e., when they are young), agents receive subsidy from the government to invest in their education. The agents use this education subsidy along with the human capital that they inherit from the previous (parental) generation to produce the human capital in this period. We assume that the agents work in the second period when they are middle aged, and retire in the third period when they are old. Let $t = 1, 2, ...$ index time. The population size of generation-t is denoted by N_t, generation-t refers to the (middle aged) agents who work at time t. At any date t, the population size is thus given by $N_{t-1} + N_t + N_{t+1}$, since at any point in time the population consists of old, middle-aged, and young agents. The population is assumed to grow at a constant rate, n.

There is a single final good in this economy and it can either be consumed in the period it is produced, or it can be stored to yield capital in the following period. The production of final good has a neo-classical constant return to scale technology which requires both physical and human capital. Thus Y_t is produced using K_t and H_t where K_t and H_t represent aggregate physical and human capital respectively at period t. But this final output production is subject to a multiplicative independent and identically distributed (i.i.d) shock λ_t at each period t. We assume that the shock λ is supported by a compact set $\left[\underline{\lambda}, \overline{\lambda}\right] \subset R_+$. This assumption is economically meaningful. We have assumed that the shock is bounded but it can be noted from our proof that this assumption has not been used to prove either the existence or the uniqueness of this steady state equilibrium. We denote the cumulated density function of λ by G and the associated probability density function by g.

Thus with per capita $h_t = \frac{H_t}{N_t}$ and defining $\overline{k_t} \equiv \frac{k_t}{h_t}$ we have the following aggregate production function

$$Y_t = \lambda_t F(K_t, H_t) = \lambda_t H_t f(\overline{k_t}) \tag{1}$$

The restrictions on f are as follows; it is positive, strictly increasing and strictly concave in its argument that is $f > 0$, $f' > 0$ and $f'' < 0$. For the purpose of analytical tractability, capital is assumed to depreciate fully between periods. Thus in this economy, the only uncertainty that we experience is the productivity shock. Note that $\overline{k_t}$ represents the factor intensity ratio which is very crucial in this analysis.

At every t, a linear tax is imposed on the income of the present middle aged agents and the tax revenue is transferred to the present young to fund the production of human capital. This constant tax rate on income is denoted by π. Moreover, since we consider an infinitely lived government, it is equivalent to saying that in this economy the government is the only source of funding for producing human capital. When the generation-t agents are young in period $t-1$, human capital or effective labor supply is produced using the human capital they have inherited from their parental generation as well as the education subsidy that is made available to them by the government. More precisely, if h_{t-1} is the per capita human capital inherited from parents and e_{t-1} is the amount of investment in education in period $t-1$ by a generation-t agent, per capita human capital produced by the middle aged agents at t is given by

$$h_t = \Phi(e_{t-1}, h_{t-1}) \tag{2}$$

We assume that the function Φ is homogenous of degree one and therefore we can write it in its intensive form as

$$h_t = h_{t-1}\Phi(\frac{e_{t-1}}{h_{t-1}}, 1) = h_{t-1}\phi(\frac{e_{t-1}}{h_{t-1}}) = h_{t-1}\phi(\overline{e}_{t-1}) \tag{3}$$

where $\overline{e}_t$ is defined as $\frac{e_t}{h_t}$. $\phi(*)$ is strictly increasing, strictly concave and satisfies the Inada conditions that is $\phi'(*) > 0$, $\phi''(*) < 0$ with $\phi'(0) = \infty$, $\phi'(\infty) = 0$.

An agent of generation-t consumes only when she is middle aged at t and old at $t+1$. She does not derive utility either from her consumption or from leisure when she is young. In the second period of her life at t, she decides her middle-aged consumption level, c_t and her saving s_t at the gross interest rate R_{t+1}. In period $t+1$, when the agent is old, she retires and consumes the return on her saving. Agents have a von-Neumann Morgenstern, time-separable utility function over consumption of final goods. This lifetime expected utility of an agent is given by

$$EU_t \equiv u(c_t) + \rho E_t u(d_{t+1}) \tag{4}$$

where c_t and d_{t+1} represent the consumptions when they are middle aged and old respectively. The parameter ρ represents the temporal discount factor which lies in between 0 and 1. E_t stands for expectation which is formed at period t for consumption in period $t+1$. We impose standard assumptions on the utility function u; it is strictly increasing and concave that means $u' > 0$ and $u'' < 0$. In the next section we formally present the problem of an agent in an economy where the government provides education subsidy for education.

2.1 Agent's problem under education subsidy

In a competitive setup, remuneration of the factors of productions are determined according to their marginal productivities. Since the final good at each period is produced using the physical and human capital available in that particular period, it can easily be shown that

$$R_t(\overline{k}_t;\lambda_t) = \lambda_t f'(\overline{k}_t) \tag{5}$$

$$w_t(\overline{k}_t;\lambda_t) = \lambda_t[f(\overline{k}_t) - \overline{k}_t f'(\overline{k}_t)] \tag{6}$$

where R_t and w_t stand for gross rate of return on physical and human capital at period t respectively. We denote the factor price ratio at t as $\Theta(\overline{k}_t;\lambda_t)$. It is clear from the above two equations that the factor price ratio can be represented as a function of the ratio of factor intensities namely $\overline{k_t}$, that is, $\Theta(\overline{k}_t;\lambda_t) = \dfrac{w_t(\overline{k}_t;\lambda_t)}{R_t(\overline{k}_t;\lambda_t)}$.

In this setup, the agents consider these factor prices as given when they maximize their lifetime expected utility. At each date t, the agents of generation- t maximize EU_t subject to their income constraint. Since w_t is the wage rate, agents receive a gross income of $w_t h_t$ when they acquire h_t amount of human capital. Note that agents do not optimally choose the education in this model. Rather they accept the amount of funds that they receive from the government as given. A generation-t agent produces human capital at $t-1$ and forms her expectation over lifetime utility for the remaining two remaining periods. This is equivalent to maximizing the expected life time utility at t given the agent has already produced human capital in $t-1$.

To meet the amount of subsidy for the current young, the government imposes a per capita proportional tax at the rate of π on the wage income of the middle aged agents. Thus at each t, a per capita total tax of $\pi w_t h_t$ is collected from each of the middle aged agents, whose population size is N_t at t. Note that since tax is imposed *after* the output is realized, the government does not face any uncertainty with respect to the tax revenue (and hence subsidy) once the output is realized. In this model, government has only two roles to play once the output is realized. At each t it collects $N_t \pi w_t h_t$ as revenue and uses this to provide a per capita subsidy of an amount e_t to those who are young at t

and has a cohort size of N_{t+1}. We assume that at each date t the budget is balanced. Thus formally

$$\begin{aligned} N_t \pi w_t h_t &= N_{t+1} e_t \; \forall t \\ &\Rightarrow \; e_t = \frac{\pi w_t h_t}{(1+n)} \; \forall t. \end{aligned} \tag{7}$$

Note that in this setup, agents take the decision of how much to consume but not how much to invest in education, since the education is funded by the government. Thus agents do not choose education optimally.

Agents maximize their lifetime expected utility (4) subject to budget constraints to make the decision about their levels of consumption and saving. Formally, the maximization problem is

$$\max_{s_t} u(c_t) + \rho E_t u(d_{t+1})$$

subject to the following constraints

$$c_t = (1-\pi) w_t h_t - s_t \tag{8}$$

$$d_{t+1} = R_{t+1} s_t. \tag{9}$$

The optimality condition is straight forward. It describes the optimum intertemporal consumption-saving decision of the agent and is characterized by the following

$$u'(c_t) = \rho E_t u'(d_{t+1}) R_{t+1}$$

Next we focus on the market clearing condition for this economy. The collective saving of the middle aged agents becomes capital stock for the next period. We assume that the market clearing condition is satisfied at each date t and thus can be written as

$$\begin{aligned} N_t s_t &= K_{t+1} \; \forall t \\ &\Rightarrow \; s_t = (1+n)\overline{k}_{t+1} h_{t+1} \; \forall t. \end{aligned} \quad (10)$$

We are now in a position to define an equilibrium of this competitive economy We call this competitive equilibrium as a Dynamic Competitive Equilibrium (DCE).

Definition 1 *A Dynamic Competitive Equilibrium (DCE) in this economy is a sequence* $\{\overline{k}_t\}_{t=0}^{\infty}$ *with initial condition* $\overline{k}_0 > 0$ *such that at every* t *an agent maximizes her expected lifetime utility* EU_t *subject to the budget constraint, final good* Y_t *is produced using physical capital* (K_t) *and human capital* (H_t)*, factors are paid according to their marginal productivities (that is, equations (5) and (6) hold), balanced budget condition (equation (7)) holds with equality and the market clearing condition (that is, equation (10)) is satisfied.*

In order to make the model analytically tractable, in the remaining analysis, we assume some specific functional forms. We assume that the instantaneous utility function u has a logarithmic representation. It can be easily checked that given this representation of the utility function, interest rates have no overall effect on consumption. We also assume that the human capital and final good production functions at each date t take the following forms respectively

$$\phi(\overline{e}_{t-1}) = A\overline{e}_{t-1}^{\alpha} \quad (11)$$

$$y_t = \lambda_t f(\overline{k_t}) = \lambda_t B \overline{k}_t^{\beta}. \quad (12)$$

From the above specifications, one can easily verify that $h_t = \Phi(e_{t-1}, h_{t-1}) = Ae_{t-1}^{\alpha} h_{t-1}^{1-\alpha} = h_{t-1}\phi(\overline{e}_{t-1})$ and $Y_t = \lambda_t F(K_t, H_t) = \lambda_t B K_t^{\beta} H_t^{1-\beta}$. Without loss of generality, we can assume that $B = 1$. Since remuneration of the factors of production are determined in a competitive market, it can easily be shown that

$$R_t = \lambda_t f'(\overline{k}_t) = \lambda_t \beta B \overline{k}_t^{\beta-1} \tag{13}$$

$$w_t = \lambda_t[f(\overline{k}_t) - \overline{k}_t f'(\overline{k}_t)] = \lambda_t(1-\beta)B\overline{k}_t^{\beta}. \tag{14}$$

Using the above two equations, we clearly have $\Theta(\overline{k}_t) = \dfrac{w_t(\overline{k}_t;\lambda_t)}{R_t(\overline{k}_t;\lambda_t)} = \frac{\beta}{1-\beta}\overline{k}_t$.

In the following theorem we present a characterization of the optimal path for effective capital in the above discussed economy. Interestingly we have been able to find out a nice and manageable expression for $\overline{k}_{t+1}$ in terms of $\overline{k}_t$ and reach to the following conclusion.

Theorem 2 *The competitive equilibrium produces a log-linear stochastic process given by*

$$\overline{k}_{t+1} = \Gamma_G(\lambda_t, \overline{k}_t) \equiv \theta_\pi \lambda_t^{1-\alpha} \overline{k}_t^{\beta(1-\alpha)} \tag{15}$$

where

$$\theta_\pi \equiv \frac{\rho(1-\pi)[(1-\beta)B]^{1-\alpha}}{(1+\rho)\pi^\alpha A(1+n)^{1-\alpha}}. \tag{16}$$

Proof. Shown in the appendix. ■

2.1.1 Invariant state under the economy with education subsidy

Conventionally in a non-stochastic case, a balanced growth path is defined as a situation where the ratio of two capital stocks is constant over time. We define an equivalent version of a balanced growth path in our model and we term it as ‘stochastic balanced growth path’ (SBGP). The concept of stochastic balanced growth path is an invariant distribution of the ratio of two types of capital stock. That means while the balanced growth path in a non-stochastic environment is a constant ratio of two stocks of capital, here in the stochastic case, it is an invariant distribution of the ratio of the two stocks as time moves. It can equivalently be characterized as a long run equilibrium in our model. Before we formally define a stochastic growth path, we first provide the standard definition of invariant distribution of the markov process $\overline{k}_t$ below.

Definition 3 *Let* $\{\overline{k}_t\}_{t\geqslant 1}$ *is defined over a probability space* $(\Omega, \mathcal{F}, P)$*. A probability measure* μ *over* $(\Omega, \mathcal{F})$ *is called an invariant distribution for* $\overline{k}_t$ *if*

$$\mu(B) = \int P(\overline{k}, B)\mu(d\overline{k}) \ \forall B \in \mathcal{B}.$$

Using the above definition, we now define our long run equilibrium as a stochastic balanced growth path.

Definition 4 *(Stochastic Balanced Growth Path or SBGP) A distribution G defines a SBGP if the invariant distribution of* $\overline{k}$ *is represented by the distribution G.*

We now show that under a very mild condition, a balanced growth path exists for the above economy. Theorem 5 below presents the results related to the existence and uniqueness of the growth path for our benchmark economy. As discussed earlier, we do not need the assumption of boundedness to show the existence of the balanced growth path.

Theorem 5 *(Existence of SBGP) Let* $E\left|\ln \varepsilon_0\right| < \infty$ *where* $\varepsilon_t \equiv \ln \theta_\pi + (1-\alpha)\ln \lambda_t$*. Then for any initial distribution of* $\overline{k}_0$*,* $\overline{k}_t$ *converges in distribution, that is, there exists a SBGP for the above economy. Further,* $\overline{k}_t \xrightarrow{d} e^\eta$ *where* $\eta \equiv \sum_{j=0}^{\infty} \rho^j \varepsilon_j$*,* $\rho \equiv \beta(1-\alpha)$ *and* $\mu(.) = P(e^\eta \in .)$ *is an invariant probability measure for* $\{\overline{k}_t\}$*. The SBGP of this economy is also unique and it has a non atomic distribution.*

Proof. The theorem can be easily proved using the proposition 1 of Athreya and Pantula (1986) (Also see Bhattacharya and Waymire (1990), Bhattacharya and Majumdar

(2007)). First of all, $\{\varepsilon_t\}$ are i.i.d and $|\rho| < 1$ where $\rho \equiv \beta(1-\alpha)$, $\varepsilon_t \equiv \ln\theta + (1-\alpha)\ln\lambda_t$. Note that a logarithmic transformation of the above recursive relation (15) can be written as

$$\ln \overline{k}_{t+1} = \ln\theta + (1-\alpha)\ln\lambda_t + \beta(1-\alpha)\ln\overline{k}_t \tag{17}$$

Next we denote $\eta_t \equiv \ln\overline{k}_t$. Then equation (17) can be written as

$$\eta_{t+1} = \rho\eta_t + \varepsilon_t, t \geqslant 0 \tag{18}$$

Note that

$$\eta_{t+1} = \rho^{t+1}\eta_0 + \sum_{j=1}^{t+1}\rho^j\varepsilon_{t+1-j}. \tag{19}$$

Since η_0, η_1, ...η_n are all determined by $\{\eta_0, \varepsilon_0, ...\varepsilon_{n-1}\}$ and ε_n is independent of ε_{n-1}, for every Borel set B, $P(\eta_{t+1} \in B \mid \{\eta_0, \eta_1, ...\eta_n\}) = Q(B - \rho\eta_t)$ where Q is the common distribution of the random variables ε_t. Therefore, $\{\eta_t\}_{t\geqslant 0}$ is a markov process in $\mathbb{R}^1$ with transition probability $p(x, B) \equiv Q(B - \rho x)$ and initial η_0. Observe that since $\{\varepsilon_t\}$ are i.i.d., $\sum_{j=1}^{t+1}\rho^j\varepsilon_{t+1-j}$ and $\sum_{j=1}^{t+1}\rho^j\varepsilon_j$ have the same distribution. Also since $\rho < 1$, $\rho^{t+1}\eta_0 \to 0$ as $t \to \infty$ almost surely. Then using Borel Cantelli Lemma, we have $\sum_{j=1}^{t+1}\rho^j\varepsilon_j \longrightarrow \eta \equiv \sum_{j=0}^{\infty}\rho^j\varepsilon_j$ almost surely when $E\left|\ln\varepsilon_0\right| < \infty$. And finally $\eta_t \equiv \ln\overline{k}_t \xrightarrow{d} \eta \Rightarrow \overline{k}_t \xrightarrow{d} e^\eta$. Also it is obvious that $\mu(.)$ is a stationary for $\left\{\overline{k}_t\right\}$. Since $\eta_{t+1} = \rho^{t+1}\eta_0 + \sum_{j=1}^{t+1}\rho^j\varepsilon_{t+1-j}$ and $\overline{\eta}_{t+1} \equiv \rho^{t+1}\eta_0 + \sum_{j=1}^{t+1}\rho^j\varepsilon_j$ have the same distribution, they both converge to the same distribution. Hence the invariant distribution for this markov process is unique. Since exponential function is a strictly increasing one, the proof follows automatically[3]. ■

Note that the above result can partially be viewed as a generalization of the results (which includes human capital) of Galor and Ryder (1989) when their model is extended to a stochastic endogenous growth setup. Stachurski (2003) has provided some conditions

[3]Note that boundedness of shock is not required to prove the existence. It is not required to prove the following three theorems either.

under which a unique and globally stable stochastic equilibrium exists in a model without human capital.

Note that using Lai and Wei (1982) one can prove that η_t has a continuous distribution provided ε_0 is not degenerate. Therefore it is straight forward to show that e^{η} has a continuous distribution. Further, using Athreya and Pantula (1986), we have the following result.

Theorem 6 *If ε_0 has a non trivial absolutely continuous component then η is fully absolutely continuous.*

Proof. See Proposition 2 of Athreya and Pantula (1986). ■

It is also easy to verify that applying the law of large number (LLN) we can present the following theorem.

Theorem 7 *For any function h such that $E\,|h(\eta)| < \infty$, $\frac{1}{n}\sum_{j=0}^{\infty} h(\eta_j) \longrightarrow E(h(\eta))$ almost surely.*

Proof. The proof can be constructed by applying law of large number (LLN) (see for e.g., Athreya and Lahiri (2006) for detail). ■

3 Presence of a perfect capital market

In this section we compare our benchmark economy with education subsidy to an economy where there exists a perfect capital market in order to see whether the same growth path can be achieved by our benchmark economy for any particular tax rate. Though in reality, the presence of a perfect capital market is rare (see Becker 1975, Kehoe and Levine 2000, Boldrin and Montes 2003, Bishnu 2009), the motivation behind this exercise is to investigate whether a benchmark economy like ours has the ability to follow the same growth path as an economy with complete market allocations. We should also mention at this point that in this paper, we were less concerned about the welfare issue; rather we focus on the growth path. More specifically, our aim is to find the tax rate at which our benchmark economy with educational subsidy and an intergenerational transfer generates the same growth path as an economy with a perfect capital market.

For this purpose, we assume that there is no government intervention in the education sector. However, there exists a perfect capital market where the young agents can borrow funds to finance their education. The agents then use this fund together with the stock of human capital available from their previous generations to produce their human capital. These young agents repay their debt when they are middle aged in the next period. Note that here agents can optimally choose their level of education by choosing the amount of funds they borrow. This is a point of diversion from our benchmark economy. Therefore, in a competitive set up agents choose their level of saving as well as the amount of capital they need to borrow for their education. Since the capital market is the *only* source of funding for education, as we shall see later, the rate of interest on borrowing for education can safely be ignored. This result is in fact not very surprising. The absence of heterogeneity in the model along with the facts that agents do not maximize utility when they are young and that agents must borrow for their survival, lead to this result.

An agent of generation t borrows an amount of e_{t-1} in period $t-1$ from a perfect capital market and invests in education. Thus when an agent receives a level of human capital h_{t-1} from his parental generation, the human capital production function can be written as

$$h_t = \phi(e_{t-1}, h_{t-1}). \tag{20}$$

Agents earn wages when they are middle aged and repay the funds they borrowed for education. Naturally we do not have any balanced budget condition here. Since a part of the collective savings of the middle aged is used to finance education for the young and the remaining amount becomes capital stock for the next period, the market clearing condition can be written as $N_t s_t = K_{t+1} + N_{t+1} e_t$ which means

$$s_t = \overline{k}_{t+1}(1+n)h_{t+1} + (1+n)e_t. \tag{21}$$

3.1 Agent's problem under perfect capital market

An agent born at $t-1$ chooses an amount of investment towards human capital so that it maximizes her expected lifetime utility. Then she chooses her optimal saving, given she has already chosen the amount of capital she wants to borrow for her education. We solve the problem backwards in two stages. First, given the level of funds she has borrowed, she maximizes her expected utility by optimally choosing her level of consumption and hence the level of saving. In the next step, we solve for the optimal level of borrowing amount by maximizing her expected indirect utility.

Thus an agent born at $t-1$ solves the following maximization problem at t given e_{t-1}:

$$\max_{s_t} EU_t \equiv u(c_t) + \rho E_t u(d_{t+1}) \tag{22}$$

subject to

$$c_t = w_t h_t - R_t e_{t-1} - s_t,$$

$$d_{t+1} = R_{t+1} s_t.$$

Assuming interior solutions, the first order condition of a generation-t agent's problem is characterized by

$$s_t : u'(c_t) = \rho E_t u'(d_{t+1}) R_{t+1}. \tag{23}$$

This optimality condition is straight forward. It simply describes the optimum intertemporal consumption-saving decision of the agent. Given the logarithmic utility function, using the above equation, we have

$$s_t = \frac{\rho\,(w_t h_t - R_t e_{t-1})}{1+\rho}. \tag{24}$$

Thus we can write the expected indirect utility v_t as

$$v_t = \log[\frac{1}{1+\rho}(w_t h_t - R_t e_{t-1})] + \rho E_t[\log R_{t+1}\{\frac{\rho}{1+\rho}(w_t h_t - R_t e_{t-1})\}].$$

In the next step, choosing e_{t-1} by maximizing the above indirect utility function gives the following relation

$$w_t\,\phi'(\overline{e}_{t-1}) = R_t \Rightarrow \phi'(\overline{e}_{t-1}) = \frac{R_t}{w_t}. \tag{25}$$

We now define the steady state in this set up as well:

Definition 8 *A Dynamic Competitive Equilibrium (DCE) in this economy is a sequence* $\{\overline{k}_t\}_{t=0}^{\infty}$ *with initial condition* $\overline{k}_0 > 0$ *such that at every* t*, an agent maximizes her expected lifetime utility* EU_t *subject to the budget constraint, final good* Y_t *is produced using physical capital* (K_t) *and human capital* (H_t)*, factor prices* R_t *and* w_t *are determined according to their marginal productivities and the market clearing condition (i.e., equation (21)) is satisfied.*

The following theorem states the path of the physical to human capital ratio in this economy.

Theorem 9 *The competitive equilibrium of this economy with a perfect capital market also produces a log-linear stochastic process given by*

$$\overline{k}_{t+1} \equiv \Gamma_P(\lambda_t, \overline{k}_t) = \theta \lambda_t^{1-\alpha} \overline{k}_t^{\beta(1-\alpha)} \tag{26}$$

where

$$\theta \equiv \left[\frac{\rho\beta(1-\alpha)(1-\beta)B}{(1+\rho)\left\{\beta+\alpha(1-\beta)\right\}(1+n)A\{\frac{A\alpha(1-\beta)}{\beta}\}^{\frac{\alpha}{1-\alpha}}} \right]^{1-\alpha}. \tag{27}$$

Proof. Shown in the appendix. ■

3.1.1 Invariant state under a perfect capital market

Theorem 10 *(Existence of SBGP) Let* $E\left|\ln \widetilde{\varepsilon}_0\right| < \infty$ *where* $\widetilde{\varepsilon}_t \equiv \ln\theta + (1-\alpha)\ln\lambda_t$ *and* $\widetilde{\eta}_t \equiv \ln \overline{k}_{t,\tau}$. *Then for any initial distribution of* $\overline{k}_0$, $\overline{k}_t$ *converges in distribution, that is, there exists a SBGP for the above economy. Further,* $\overline{k}_t \overset{d}{\longrightarrow} e^{\widetilde{\eta}}$ *where* $\widetilde{\eta} \equiv \sum\limits_{j=0}^{\infty}\rho^j\widetilde{\varepsilon}_j$, $\rho \equiv \beta(1-\alpha)$ *and* $\mu(.) = P(e^{\widetilde{\eta}} \in .)$ *is an invariant probability measure for* $\left\{\overline{k}_t\right\}$. *The SBGP of this economy is also unique and it has a non atomic distribution.*

Proof. The proof is similar to the proof of Theorem 5. ■

The following two theorems are similar in spirit to theorems 6 and 7.

Theorem 11 *If* $\widetilde{\varepsilon}_0$ *has a non trivial absolutely continuous component then* $\widetilde{\eta}$ *is fully absolutely continuous.*

Theorem 12 *For any function* h *such that* $E\left|h(\widetilde{\eta})\right| < \infty$, $\frac{1}{n}\sum\limits_{j=0}^{\infty}h(\widetilde{\eta}_j) \longrightarrow E(h(\widetilde{\eta}))$ *almost surely.*

4 Observations

Proposition 13 *There exists a unique tax rate* $\pi = \tau$ *for which the economy with education subsidy shares the same path of a first best economy.*

Proof. Let us define the term $V(\pi) \equiv \frac{\theta_\pi}{\theta}$, $\theta > 0$. It can easily be checked that $V(\pi)$ exists since $\theta \neq 0$. Note that $\lim\limits_{\pi\to 0} V(\pi) \to \infty$ and $\lim\limits_{\pi\to 1} V(\pi) \to 0$. Since we have assumed that $\pi \in (0,1)$, by continuity of $V(\pi)$, there exists a $\pi = \tau$ such that $V(\tau) = 1$. Further, it can also be checked that $V'(\pi) < 0$, i.e., $V(\pi)$ is strictly monotonically decreasing in π. Hence the proof. ■

For the purpose of notational convenience, from this point onwards we denote the path of $\overline{k}_{t+1}$ in an economy with a perfect capital market by

$$\begin{aligned} \overline{k}_{t+1,\tau} &= \Gamma_P(\lambda_t, \overline{k}_{t,\tau}) = \theta_\tau \lambda_t^{1-\alpha} \overline{k}_{t,\tau}^{\beta(1-\alpha)} \text{ where} \\ \theta_\tau &= \left[\frac{\rho\beta(1-\alpha)(1-\beta)B}{(1+\rho)\left\{\beta + \alpha(1-\beta)\right\}(1+n)A\{\frac{A\alpha(1-\beta)}{\beta}\}^{\frac{\alpha}{1-\alpha}}} \right]^{1-\alpha}. \end{aligned} \tag{28}$$

Thus the path of $\overline{k}_{t+1}$ associated with the subsidized economy and an economy with a perfect capital market are denoted by $\overline{k}_{t+1,\pi}$ and $\overline{k}_{t+1,\tau}$ respectively. We now focus on the steady state of these two paths. We start by determining the transition function. We should mention here that to prove the following lemma and the proposition, we found Hauenschild (2002) very helpful. We have also provided an alternative proof of our main claim but for the proof of stochastic dominance, we follow the method of Hauenschild (2002).

The transition function of $\left\{\overline{k}_{t,\pi}\right\}$ and $\left\{\overline{k}_{t,\tau}\right\}$ is denoted by $P_i(\overline{k}_t, B)$, $i = \pi, \tau$ respectively where

$$P_i(\overline{k}_t, B) = \int \underline{\mathbf{1}}_{B_{\overline{k}_{t,i}}}(\lambda_t) dG(\lambda_t)$$

$\forall \overline{k}_t > 0$ and $\forall B \in \mathcal{B}_+$, Borel-$\sigma$-algebra over $\mathcal{R}_+$. While $B_{\overline{k}_{t,i}} \equiv \left\{\lambda_t : \Gamma_i(\lambda_t, \overline{k}_{t,i}) \in B\right\}$, $\underline{\mathbf{1}}_{B_{\overline{k}_{t,i}}}(\lambda_t)$ represents an indicator function, that is.

$$\underline{\mathbf{1}}_{B_{\overline{k}_{t,i}}}(\lambda_t) = \begin{cases} 1 \text{ if } \lambda_t \in B_{\overline{k}_{t,i}} \\ 0 \quad \text{otherwise.} \end{cases}.$$

Lemma 14 *Let both the processes $\overline{k}_{t+1,\pi}$ and $\overline{k}_{t+1,\tau}$ start from the same initial value of $\overline{k}_{t=0}$, say $\overline{k}_0 > 0$. Then $\overline{k}_{t+1,\pi} > \overline{k}_{t+1,\tau}$ for all t on almost all the paths whenever the tax* *rate* $\pi < \tau$.

Proof. We need to show that $P\{\omega \in \Omega : \overline{k}_{t+1,\pi}(\omega) > \overline{k}_{t+1,\tau}(\omega)\} = 1$. From $t = 0$ to $t = 1$ it can easily be checked that

$$\overline{k}_{1,\pi} = \Gamma_\pi(\lambda_{t=0}, \overline{k}_{t=0,\pi}) = \theta_\pi \lambda_{t=0}^{1-\alpha} \overline{k}_{t=0,\pi}^{\beta(1-\alpha)} > \theta_\tau \lambda_{t=0}^{1-\alpha} \overline{k}_{t=0,\tau}^{\beta(1-\alpha)} = \Gamma_\tau(\lambda_{t=0}, \overline{k}_{t=0,\tau}) = \overline{k}_{1,\tau}$$

since $\theta_\pi > \theta_\tau$ and $t = 0$. Thus what remains to be shown is that our claim holds from any t to $(t+1)$. Assuming that $\overline{k}_{t,\pi} > \overline{k}_{t,\tau}$, it is easy to show that $\overline{k}_{t+1,\pi} > \overline{k}_{t+1,\tau}$ holds for any t when $\pi < \tau$. Therefore by induction, $\overline{k}_{t+1,\pi} > \overline{k}_{t+1,\tau}$ holds for any t. Hence the proof. ■

Definition 15 *Let* μ *and* ν *denote probability distributions with support* $T \subset R+$ *and let the corresponding cumulated density functions be* F *and* G *respectively. Then* ν *is said to dominate* μ *in the sense of First Order Stochastic Dominance (FOSD) if the inequality* $G(t) \leqslant F(t)$ *holds for all* t *and if* $G(t_0) < F(t_0)$ *for at least one* $t_0 \in T$.

The above definition indicates that in the sense of FOSD, the probability distribution of a SBGP associated with higher tax rate in an economy dominates the corresponding distribution of SBGP associated with relatively lower tax rate.

Proposition 16 *An SBGP with tax rate* $\pi < \tau$ *First order Stochastically Dominates an SBGP with tax rate* τ *and thus* $E(\overline{k}_\pi) > E(\overline{k}_\tau)$.

Proof. We start the proof with $t = t_0 + 1$. Notice that for any $x > 0$

$$F_{t_0+1,\pi}(x) = \mu_{t_0+1,\pi}((0,x]) = \int P_\pi(\overline{k}_{t=0,\pi}, (0,x])\mu_{t_0,\pi}(d\overline{k}_{t=0,\pi})$$

and

$$F_{t_0+1,\tau}(x) = \mu_{t_0+1,\tau}((0,x]) = \int P_\tau(\overline{k}_{t=0,\tau}, (0,x])\mu_{t_0,\tau}(d\overline{k}_{t=0,\tau}).$$

Since the two economies share the same starting point, to have

$$F_{t_0+1,\pi}(x) < F_{t_0+1,\tau}(x) \tag{29}$$

it is sufficient to show that $P_\pi(\overline{k}_{t=0,\pi}, (0,x]) < P_\tau(\overline{k}_{t=0,\tau}, (0,x])$. Since from Lemma 14, $\overline{k}_{t+1,\pi} > \overline{k}_{t+1,\tau}$ holds for any t, for $t = t_0$, we have $\overline{k}_{t_0+1,\pi} > \overline{k}_{t_0+1,\tau}$ for all possible $\lambda_{t=0}$, that is, for all λ_{t_0}. This in turn implies that for any $\overline{k}_{t=0,i}$

$$\left\{\lambda_{t_0} : \overline{k}_{t+1,\pi} \in (0,x]\right\} \subset \left\{\lambda_{t_0} : \overline{k}_{t+1,\tau} \in (0,x]\right\}. \tag{30}$$

Note that (30) implies $P_\pi(\overline{k}_{t=0,\pi}, (0,x]) < P_\tau(\overline{k}_{t=0,\tau}, (0,x])$ and thus the result holds for $t = t_0$. In the next step we show that $F_{t+1,\pi}(x) \leqslant F_{t+1,\tau}(x)$ holds for any t. Without loss of generality we assume that $F_{t,\pi}(x) \leqslant F_{t,\tau}(x)$ holds at t. Given this, we prove that $F_{t+1,\pi}(x) \leqslant F_{t+1,\tau}(x)$ for any x. Note that for any t and $x > 0$ and $\overline{k}_t$ we have

$$F_{t+1,\pi}(x) = \mu_{t+1,\pi}((0,x]) = \int P_\pi(\overline{k}_t(0,x])\mu_{t,\pi}(d\overline{k}_{t,\pi})$$

and

$$F_{t+1,\tau}(x) = \mu_{t+1,\tau}((0,x]) = \int P_\tau((\overline{k}_t, (0,x])\mu_{t,\tau}(d\overline{k}_{t,\tau}).$$

We do not use subscript π or τ when the corresponding values are same. First we show that

$$F_{t+1,\pi}(x) = \mu_{t+1,\pi}((0,x]) = \int P_\pi(\overline{k}_t, (0,x])\mu_{t,\pi}(d\overline{k}_{t,\pi}) \leqslant \int P_\pi(\overline{k}_t, (0,x])\mu_{t,\tau}(d\overline{k}_{t,\tau}).$$

Let $f_{t,\pi}$ and $f_{t,\tau}$ are the associated probability density functions associated with $\mu_{t,\pi}$ and

$\mu_{t,\tau}$. Thus

$$\begin{aligned} & \int P_\pi(\overline{k}_t,(0,x])\mu_{t,\pi}(d\overline{k}_{t,\pi}) - \int P_\pi((\overline{k}_t,(0,x])\mu_{t,\tau}(d\overline{k}_{t,\tau}) \\ = & \int P_\pi((\overline{k}_t,(0,x])(f_{t,\pi}(\overline{k}_t) - f_{t,\tau}(\overline{k}_t))d\overline{k}_t. \end{aligned} \tag{31}$$

At this stage we should find out $P_\pi((\overline{k}_t,(0,x])$ in our set up. Since we have assumed that the shock is bounded, it can be shown that

$$P_\pi((\overline{k}_t,(0,x]) = \begin{cases} 1 & \text{if } \overline{k}_t \leqslant \overline{k}_x^{(1)} \\ G\left(\dfrac{x^{\frac{1}{1-\alpha}}}{\theta_\pi^{\frac{1}{1-\alpha}}\overline{k}_t^{\beta}}\right) & \text{if } \overline{k}_t \in (\overline{k}_x^{(1)},\overline{k}_x^{(2)}) \\ 0 & \text{if } \overline{k}_t \leqslant \overline{k}_x^{(2)} \end{cases}$$

where

$$\overline{k}_x^{(1)} = \left(\frac{x}{\theta_\pi\overline{\lambda}^{1-\alpha}}\right)^{\frac{1}{\beta(1-\alpha)}} \text{ and } \overline{k}_x^{(2)} = \left(\frac{x}{\theta_\pi\underline{\lambda}^{1-\alpha}}\right)^{\frac{1}{\beta(1-\alpha)}}.$$

Thus (31) can be written as

$$\begin{aligned} & \int P_\pi((\overline{k}_t,(0,x])(f_{t,\pi}(\overline{k}_t) - f_{t,\tau}(\overline{k}_t))d\overline{k}_t \\ = & \int_{(0,\overline{k}_{x,\pi}^{(1)}]} 1.(f_{t,\pi}(\overline{k}_t) - f_{t,\tau}(\overline{k}_t))d\overline{k}_t + \int_{(\overline{k}_{x,\pi}^{(1)},\overline{k}_{x,\pi}^{(2)})} G\left(\frac{x^{\frac{1}{1-\alpha}}}{\theta_\pi^{\frac{1}{1-\alpha}}\overline{k}_t^{\beta}}\right)(f_{t,\pi}(\overline{k}_t) - f_{t,\tau}(\overline{k}_t))d\overline{k}_t \\ = & F_{t,\pi}(\overline{k}_x^{(1)}) - F_{t,\tau}(\overline{k}_x^{(1)}) + G\left(\frac{x^{\frac{1}{1-\alpha}}}{\theta_\pi^{\frac{1}{1-\alpha}}\overline{k}_t^{\beta}}\right)\left(F_{t,\pi}(\overline{k}_t) - F_{t,\tau}(\overline{k}_t)\right)\Big|_{(\overline{k}_{x,\pi}^{(1)},\overline{k}_{x,\pi}^{(2)})} \qquad (32) \\ & - \int_{(\overline{k}_{x,\pi}^{(1)},\overline{k}_{x,\pi}^{(2)})} \frac{\partial G}{\partial \overline{k}_t}\left(\frac{x^{\frac{1}{1-\alpha}}}{\theta_\pi^{\frac{1}{1-\alpha}}\overline{k}_t^{\beta}}\right)(F_{t,\pi}(\overline{k}_t) - F_{t,\tau}(\overline{k}_t))d\overline{k}_t. \qquad (33) \end{aligned}$$

Given the assumption, the above equation is enough to prove the first part. In the next step we show that

$$\int P_\pi(\overline{k}_t,(0,x])\mu_{t,\tau}(d\overline{k}_{t,\tau}) \leqslant \int P_\tau(\overline{k}_t,(0,x])\mu_{t,\tau}(d\overline{k}_{t,\tau}).$$

The above inequality can be written as

$$\int \left\{P_\pi(\overline{k}_t,(0,x]) - P_\tau(\overline{k}_t,(0,x])\right\}\mu_{t,\tau}(d\overline{k}_{t,\tau}) \leqslant 0. \tag{34}$$

Since the outer integral is same, to prove (34) it is enough to show that $P_\pi(\overline{k}_t,(0,x]) \leqslant P_\tau(\overline{k}_t,(0,x])$. It can be easily verified using an induction similar to the one used at the beginning of the proof that the above relation holds at any t. Hence the proof. ■

If an SBGP with tax rate π First order Stochastically Dominates a SBGP with tax rate τ, it implies that $E(\overline{k}_\pi) > E(\overline{k}_\tau)$ holds but the converse is not necessarily true. Thus, showing that $E(\overline{k}_\pi) > E(\overline{k}_\tau)$ holds is not enough to claim that a SBGP with tax rate π First order Stochastically Dominates a SBGP with tax rate τ. We now provide an alternative proof to establish that $E(\overline{k}_\pi) > E(\overline{k}_\tau)$ holds.

An alternative proof of $E(\overline{k}_\pi) > E(\overline{k}_\tau)$ when $\pi < \tau$:

Proof. From theorems 5 and 12 we know that for both $\sigma_j = \eta_j, \widetilde{\eta}_j$ and for any function h such that $E\left|h(\sigma)\right| < \infty$, $\frac{1}{n}\sum_{j=0}^{\infty} h(\sigma_j) \longrightarrow E(h(\sigma))$ almost surely. Note that like **(19), (28)** can be written as

$$\widetilde{\eta}_{t+1} = \rho\widetilde{\eta}_t + \widetilde{\varepsilon}_t, t \geqslant 0. \tag{35}$$

We next define $\Delta \equiv \varepsilon_t - \widetilde{\varepsilon}_t$. We do not put any time subscript for Δ since it can be easily checked that Δ is time invariant. We find that

$$\Delta = \log\frac{\theta_\pi}{\theta_\tau}.$$

It can also be checked that given our specification, $\Delta > 0$. Define $\delta E \equiv E\left(e^{\widetilde{\eta}}\right) - E\left(e^{\eta}\right)$ and we show that $\delta E < 0$. Note that

$$\begin{aligned}
\delta E &= E\left(e^{\sum_{j=0}^{\infty}\rho^j \widetilde{\varepsilon}_j}\right) - E\left(e^{\sum_{j=0}^{\infty}\rho^j \varepsilon_j}\right) \\
&\Rightarrow \delta E = E\left(e^{\sum_{j=0}^{\infty}\rho^j (\varepsilon_j - \Delta)}\right) - E\left(e^{\sum_{j=0}^{\infty}\rho^j \varepsilon_j}\right) \\
&= E\left(e^{\sum_{j=0}^{\infty}\rho^j (\varepsilon_j - \Delta)} - e^{\sum_{j=0}^{\infty}\rho^j \varepsilon_j}\right) \\
&= \left(\frac{1}{e^{\frac{\Delta}{1-\rho}}} - 1\right) E\left(e^{\sum_{j=0}^{\infty}\rho^j \varepsilon_j}\right) < 0.
\end{aligned}$$

Thus $E(\overline{k}_\pi) > E(\overline{k}_\tau)$. ■

5 Conclusion

In this paper, we consider an economy where agents receive a subsidy from the government to fund their education. This subsidy is made available through a proportional tax scheme that is imposed on the working generation. The tax revenue is used to fund this education subsidy in a way so that the budget is balanced. The production of the final good uses both human and physical capital. Furthermore, this production process is subject to a shock. This gives rise to a stochastic environment. We define a steady state equilibrium in this stochastic setup and discuss the concept of a balanced growth path in this framework. We show that under reasonable assumptions this steady state exists and is unique. We further show that for a particular tax rate, this benchmark economy shares the same steady state distribution as an economy with a perfect capital market and thus first best. Moreover, it has been shown that for a tax rate lower than this particular tax rate, the steady state of our benchmark economy first order stochastically dominates the steady state of the latter economy where a perfect capital market is present.

References

[1] Amir, R., A new look at optimal growth under uncertainty, Journal of. Economic. Dynamics and. Control 22 (1997), 67–86.

[2] Athreya, K. B., Lahiri, S., Measure theory and probability theory, Springer. 2006

[3] Athreya, K. B., Pantula, Sastry G., Mixing Properties of Harris Chains and Autoregressive Processes, Journal of Applied Probability 23 (1986), 880-892

[4] Becker, G., S., University of Chicago Press, Chicago (1975)

[5] Bhattacharya, R., Majumdar, M., On a class of stable random dynamical systems: theory and applications, Journal of Economic Theory 96 (2001), 208-291

[6] Bhattacharya, R., Majumdar, M., Random dynamical systems, Cambridge University Press, (2007)

[7] Bhattacharya, R., Waymire, E. C., Stochastic Processes with Applications (1990), Wiley, New York

[8] Bishnu, M., Government Intervention in Education in the Presence of Consumption Externalities, working paper

[9] Boldrin, M., Public Education and Capital Accumulation" (C.M.S.E.M.S. Discussion Paper 1017 Northwestern University) (1992)

[10] Boldrin, M., Montes, A., Assessing the efficiency of public education and pensions, Unpublished, 2004

[11] Boldrin, M., Montes, A., The intergenerational state education and pensions, Review of Economic Studies 72 (2005) 651–664.

[12] Boldrin, M., Rustichini A., Political Equilibria with Social Security". Review of Economic Dynamics, 3 (2000) 41-78

[13] Brock, W. A., Majumdar, M., Global asymptotic stability results for multisector models of optimal growth under uncertainty when future utilities are discounted, Journal of Economic Theory 18 (1978), 225–243

[14] Cass, D., Optimum growth in an aggregative model of capital accumulation, Review of Economic Studies 32 (1965), 233–240

[15] De la Croix, D., Michel, P., A Theory of Economic Growth: Dynamics and Policy in Overlapping Generations, Cambridge University Press, Cambridge, 2002.

[16] Donaldson, J. B., Mehra, R., Stochastic growth with correlated production shocks, Journal of Economic Theory 29 (1983), 282–312.

[17] Docquier, F., Paddison, O., Pestieau, P.,Optimal accumulation in an endogenous growth setting with human capital. Journal of Economic Theory 134, pp 361-378

[18] Galor, O., Ryder, H. E., Existence, uniqueness and stability of equilibrium in an overlapping generations model with productive capital. Journal of Economic Theory 49 (1989), 360–375.

[19] Hauenschild, N., Capital accumulation in a stochastic overlapping generation model with social security. Journal of Economic Theory 106 (2002), pp 201-216

[20] Hopenhayn, H. A., Prescott, E. C., Stochastic monotonicity and stationary distributions for dynamic economies. Econometrica 60 (1992),1387–1406

[21] Kehoe, T J., Levine, D. K., Liquidity Constrained Markets versus Debt Constrained Markets, Econometrica 69 (2000),575-598.

[22] Koopmans, T., On the concept of optimal economic growth, Pontif. Acad. Sci. 28 (1965), 225–300.

[23] Kotlikoff, L. J., Persson, T., Svensson, L. E. O., Social Contracts as Assets: A Possible Solution to the Ttme Consistency Problem". American Economic Review 78 (1988) 662-677.

[24] Lai, T. L., Wei, C.A., Least squares estimates in stochastic regression models with applications to identification and control of dynamic systems. Annals of Statistics (1982) 10, 154-166.

[25] Laitner, J., The steady states of a stochastic decentralized growth model, Journal of Economic Theory 24 (1981), 377–392.

[26] Majumdar, M., Zilcha, I., Optimal growth in a stochastic environment: Some sensitivity and turnpike results, Journal of Economic Theory 43 (1987), 116–133.

[27] Mirman, L. J., On the existence of steady state measures for one sector growth models with uncertain technology, International Economic Review 13 (1972), 271–286.

[28] Mirman, L. J., The steady state behavior of a class of one sector growth models with uncertain technology, Journal of Economic Theory 6 (1973), 219–242.

[29] Mirman L. J., Zilcha, I., On optimal growth under uncertainty, Journal of Economic Theory 11 (1975), 329–339.

[30] Nishimura, K., Stachurski, J., Stability of stochastic optimal growth models: a new approach, Journal of Economic Theory 122 (2005) 100–118.

[31] Ramsey, F. P., A mathematical theory of saving, Economic Journal 38 (1928), 543–559.

[32] Razin A., Yahav, J. A., On stochastic models of economic growth, International Economic Review 20 (1979), 599–604.

[33] Stachurski, J., Economic dynamical systems with multiplicative noise. Journal of Mathematical Economics 39 (2003), 135-152.

[34] Stachurski, J., Stochastic Optimal Growth with Unbounded Shock, Journal of Economic Theory 106, 40–65 (2002).

[35] Stokey, N. L., Lucas, R. E., Prescott E. C., Recursive Methods in Economic Dynamics, (1989) Harvard University Press, MA.

[36] Wang, Y., Stationary equilibria in an overlapping generations economy with stochastic production. Journal of Economic Theory 61 (1993), 423–435.

Appendix

1. Obtaining Equation (15) in Theorem 2:

Given logarithmic form of utility function and using the optimality condition, we have

$$\begin{aligned}\frac{1}{c_t} &= \rho E\{u'(d_{t+1})R_{t+1}\} \\ &\Rightarrow \frac{1}{c_t} = \frac{\rho}{s_t} \Rightarrow \rho[(1-\pi)w_t h_t - s_t] = s_t \\ &\Rightarrow s_t = \frac{\rho(1-\pi)}{1+\rho} w_t h_t. \end{aligned} \tag{36}$$

Now using market clearing equation (10), we have

$$\begin{aligned} s_t &= (1+n)\bar{k}_{t+1}h_{t+1} \Rightarrow \frac{\rho(1-\pi)}{1+\rho} w_t h_t = (1+n)\bar{k}_{t+1}h_{t+1} \\ &\Rightarrow \bar{k}_{t+1} = \frac{\frac{\rho(1-\pi)}{1+\rho} w_t h_t}{(1+n)h_{t+1}} = \frac{\rho(1-\pi)w_t h_t}{(1+\rho)(1+n)h_{t+1}}. \end{aligned} \tag{37}$$

Given $\phi(\bar{e}_{t-1}) = A\bar{e}_{t-1}^{\alpha}$ in equation (11), and $h_t = h_{t-1}\phi(\bar{e}_{t-1})$, we have

$$\frac{h_{t+1}}{h_t} = \phi(\bar{e}_t) = A\bar{e}_t^{\alpha}.$$

Thus using the above equation and equation (37), we have

$$\begin{aligned} \bar{k}_{t+1} &= \frac{\rho(1-\pi)w_t h_t}{(1+\rho)(1+n)h_{t+1}} = \frac{\rho(1-\pi)w_t}{(1+\rho)(1+n)\frac{h_{t+1}}{h_t}} \\ &\Rightarrow \bar{k}_{t+1} = \frac{\rho(1-\pi)w_t}{(1+\rho)(1+n)A\bar{e}_t^{\alpha}}. \end{aligned} \tag{38}$$

Now from the balanced budget condition (7),

$$e_t = \frac{\pi w_t h_t}{(1+n)} \Rightarrow \frac{e_t}{h_t} \equiv \bar{e}_t = \frac{\pi w_t}{(1+n)}. \tag{39}$$

From the last two equations, we have

$$\begin{aligned}\overline{k}_{t+1} &= \frac{\rho(1-\pi)w_t}{(1+\rho)(1+n)A\overline{e}_t^{\alpha}} = \frac{\rho(1-\pi)w_t}{(1+\rho)(1+n)A[\frac{\pi w_t}{(1+n)}]^{\alpha}} \\ &\Rightarrow \; \overline{k}_{t+1} = \frac{\rho(1-\pi)w_t^{1-\alpha}}{(1+\rho)(1+n)^{1-\alpha}A\pi^{\alpha}}. \qquad (40)\end{aligned}$$

Since from (14), $w_t = \lambda_t(1-\beta)B\overline{k}_t^{\beta}$, we have

$$\begin{aligned}\overline{k}_{t+1} &= \frac{\rho(1-\pi)w_t^{1-\alpha}}{(1+\rho)(1+n)^{1-\alpha}A\pi^{\alpha}} = \frac{\rho(1-\pi)\left\{\lambda_t(1-\beta)B\overline{k}_t^{\beta}\right\}^{1-\alpha}}{(1+\rho)(1+n)^{1-\alpha}A\pi^{\alpha}} \\ &\Rightarrow \; \overline{k}_{t+1} = \frac{\rho(1-\pi)\left\{(1-\beta)B\right\}^{1-\alpha}}{(1+\rho)\pi^{\alpha}A(1+n)^{1-\alpha}}\lambda_t^{1-\alpha}\overline{k}_t^{\beta(1-\alpha)}\end{aligned}$$

which is equation (15). Hence the proof.

2. Obtaining Equation (26) in Theorem 9:

From the market clearing condition under perfect capital market, namely equation (21), we have

$$\begin{aligned}s_t &= \overline{k}_{t+1}(1+n)h_{t+1} + (1+n)e_t \\ \overline{k}_{t+1} &= \frac{s_t}{(1+n)h_{t+1}} - \frac{e_t}{h_{t+1}} = \frac{s_t}{(1+n)h_{t+1}} - \frac{e_t}{Ah_t^{1-\alpha}e_t^{\alpha}} \\ &\Rightarrow \; \overline{k}_{t+1} = \frac{s_t}{(1+n)h_{t+1}} - \frac{\overline{e}_t^{1-\alpha}}{A}. \qquad (41)\end{aligned}$$

From equation (24), we have

$$\begin{aligned} s_t &= \frac{\rho\left(w_t h_t - R_t e_{t-1}\right)}{1+\rho} \\ &\Rightarrow\ s_t = \frac{\rho\left[w_t h_t - w_t \phi'(\overline{e}_{t-1}) e_{t-1}\right]}{1+\rho} \text{ using equation (25)} \\ &\Rightarrow\ s_t = \frac{\rho}{1+\rho} w_t h_t \left[1 - \frac{\phi'(\overline{e}_{t-1}) e_{t-1}}{\phi(\overline{e}_{t-1}) h_{t-1}}\right] \\ &\Rightarrow\ s_t = \frac{\rho}{1+\rho} w_t h_t \left[1 - \frac{\phi'(\overline{e}_{t-1}) \overline{e}_{t-1}}{\phi(\overline{e}_{t-1})}\right] = \frac{\rho}{1+\rho} w_t h_t \left[1-\alpha\right]. \end{aligned} \tag{42}$$

Thus using equations (41) and (42), we have

$$\begin{aligned} \overline{k}_{t+1} &= \frac{\rho\left[1-\alpha\right] w_t h_t}{(1+\rho)(1+n) h_{t+1}} - \frac{\overline{e}_t^{1-\alpha}}{A} \\ &\Rightarrow\ \overline{k}_{t+1} = \frac{\rho\left[1-\alpha\right] w_t}{(1+\rho)(1+n)\frac{h_{t+1}}{h_t}} - \frac{\overline{e}_t^{1-\alpha}}{A} \\ &\Rightarrow\ \overline{k}_{t+1} = \frac{\rho\left[1-\alpha\right] w_t}{(1+\rho)(1+n)\phi(\overline{e}_t)} - \frac{\overline{e}_t^{1-\alpha}}{A}. \end{aligned} \tag{43}$$

But from equation (25), we have

$$\begin{aligned} \phi'(\overline{e}_{t-1}) &= \frac{R_t}{w_t} = \frac{\beta}{1-\beta} \overline{k}_t^{-1} \\ &\Rightarrow\ A\alpha \overline{e}_{t-1}^{\alpha-1} = \frac{\beta}{1-\beta} \overline{k}_t^{-1} \Rightarrow \overline{e}_{t-1}^{\alpha-1} = \frac{\beta}{A\alpha(1-\beta)} \overline{k}_t^{-1} \\ &\Rightarrow\ \frac{1}{\overline{e}_t^{1-\alpha}} = \frac{\beta}{A\alpha(1-\beta)\overline{k}_{t+1}}. \end{aligned} \tag{44}$$

Thus from (43), we have

$$\begin{aligned}
\overline{k}_{t+1} &= \frac{\rho\,[1-\alpha]\,w_t}{(1+\rho)(1+n)\phi(\overline{e}_t)} - \frac{A\alpha(1-\beta)\overline{k}_{t+1}}{\beta A} \\
&\Rightarrow\ \overline{k}_{t+1} + \frac{A\alpha(1-\beta)\overline{k}_{t+1}}{\beta A} = \frac{\rho\,[1-\alpha]\,w_t}{(1+\rho)(1+n)\phi(\overline{e}_t)} \\
&\Rightarrow\ \overline{k}_{t+1} = \frac{\rho\beta(1-\alpha)w_t}{(1+\rho)\left\{\beta+\alpha(1-\beta)\right\}(1+n)\phi(\overline{e}_t)} \\
&\Rightarrow\ \overline{k}_{t+1} = \frac{\rho\beta(1-\alpha)\lambda_t(1-\beta)B\overline{k}_t^{\beta}}{(1+\rho)\left\{\beta+\alpha(1-\beta)\right\}(1+n)\phi(\overline{e}_t)} \\
&\Rightarrow\ \overline{k}_{t+1} = \frac{\beta(1-\alpha)\lambda_t(1-\beta)B\overline{k}_t^{\beta}}{(1+\rho)\left\{\beta-\alpha(1-\beta)\right\}(1+n)A\overline{e}_t^{\alpha}} \\
&\Rightarrow\ \overline{k}_{t+1} = \frac{\rho\beta(1-\alpha)\lambda_t(1-\beta)B\overline{k}_t^{\beta}}{(1+\rho)\left\{\beta+\alpha(1-\beta)\right\}(1+n)A\{\frac{A\alpha(1-\beta)}{\beta}\overline{k}_{t+1}\}^{\frac{\alpha}{1-\alpha}}} \\
&\Rightarrow\ \overline{k}_{t+1}^{\frac{1}{1-\alpha}} = \frac{\rho\beta(1-\alpha)\lambda_t(1-\beta)B\overline{k}_t^{\beta}}{(1+\rho)\left\{\beta+\alpha(1-\beta)\right\}(1+n)A\{\frac{A\alpha(1-\beta)}{\beta}\}^{\frac{\alpha}{1-\alpha}}} \\
&\Rightarrow\ \overline{k}_{t+1} = \left[\frac{\rho\beta(1-\alpha)(1-\beta)B}{(1+\rho)\left\{\beta+\alpha(1-\beta)\right\}(1+n)A\{\frac{A\alpha(1-\beta)}{\beta}\}^{\frac{\alpha}{1-\alpha}}}\right]^{1-\alpha} \lambda_t^{1-\alpha}\overline{k}_t^{\beta(1-\alpha)}
\end{aligned}$$

which is equation (26). Hence the proof.