International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 3 | Issue 1 | ISSN : 2456-3307

Predicting Student's Performance using Data on Internet Technology Usage Behavioural Patterns

Esther Khakata*, Vincent Omwenga, Simon Msanjila
Faculty of Information Technology, Strathmore University, Nairobi, Kenya

ABSTRACT

Predicting student performance is becoming increasingly important to the learners and other stakeholders. This is due to the central role that prediction plays in planning for resources in a constrained learning environment. Availability of the large educational dataset has promoted the emergence of various analytical approaches aimed at providing accurate prediction. Most of these approaches have relied on data that measure specific outcome rather than tracking the process that leads to a particular outcome. This paper proposes a Stochastic Differential Equation predictive model that evaluates in an analytic way the process leading to a certain level in student performance. The model will use data based on the behavioural patterns in the use of Internet Technology.

Keywords: Internet technology, Student Performance, Stochastic Differential Equation (SDE), Predictive Model, Behavioural Patterns

I. INTRODUCTION

The impact of Internet Technology has been beyond imaginable limits across the various spheres of life. For instance, in the areas of business, education and governance the impact has been unparalleled. It is observed that the impact of Internet Technology has been precipitated by the rapidly changing technology advancements that have promoted adoptability and ease of use of the technology. Others aspects that have also accelerated the assimilation of Internet Technology include its flexibility, cost effectiveness, availability and dynamism. [1] notes that the migration from Web 1.0 which was generally static to a more dynamic Web 2.0 has seen tremendous growth on internet uptake due to the ability to manage internet resources by reading, writing, modifying and updating content online in a dynamic way.

[2] notes that the Internet and the World Wide Web (WWW) have significantly changed the teaching and learning process in higher learning institutions. [3] supports this argument when they note that by 2020 educational institutions will be available to majority who would have otherwise missed out by leveraging on the power of the internet. The aspects of education that will be impacted most include accessibility of the content by learners and improved quality of learning. However, in order to fully harness the impact of Internet Technology, understanding the context of its usage, like any other technology will be vital [3]. This is because impact is utility driven and is contextualized in an environment where it is used. This paper provides an analysis of the student behavioural pattern on the usage of Internet Technology in their learning process. The behavioural patterns analysis is anchored on the utility aspects of Internet Technology by borrowing from Cobb-Douglas production theorem.

Student performance has been described to be a measure of student performance based on learning assessments [5]. These approaches are minimalistic since they consider student performance to be an event which is dependent on specific instantaneous times when an assessment is conducted. In the context of this paper we consider student performance as a process whose assessment should be the sum-total of the competencies acquired over time that enables the student to 1) make a correct choice based on the available alternatives, 2) execute an action with minimal costs and 3) optimize the available resources toward the attainment of the desired outcome.

In the learning environment, Internet Technology has been used as an infrastructural resource upon which educational platforms and learning resources are embedded. Some of the educational platforms and resources that have been built to ride on Internet Technology include the eLearning architectures like Moodle, ATutor, Claroline, Dokeos, Blackboard, LMS-Qstutor among others. A number of learning resources have been developed and deployed on these educational platforms for example Massive Open Online Courses (MOOCs).

It is evident that without the Internet Technology, these educational platforms and resources might not be effectively utilized [6]. Moreover, limited knowledge of Internet Technology implies limited usage of the educational resources that are supplanted on it. Studies have shown existence of a link between the level of usage of the educational resources and Internet Technology. [7] and [8] have shown that to optimally realize the benefits of using educational resources available online, aspects like knowledge, effort and costs surrounding Internet Technology needs to be addressed.

Due to the emergence of a variety of educational platforms and resources, students in a learning ecosystem are bound to use a variety of them. Studies have been conducted to determine how a particular platform or resource influences student results [9]. We consider this to be limiting since different platforms and resources exist side-by-side in most learning environments. We therefore propose an approach that looks at generic factors that may interplay on foundations upon which these educational resources are built, the Internet Technology. The proposed approach will analyse how the utility of the Internet Technology from a resource usage behavioural pattern by the student may influence the use of these education resources hence affecting student performance.

II. METHODOLOGY

Internet Technology usage by students is dependent on the behavioural intentions as can be explained by the Technology Acceptance Model (TAM). According to the theory of Reasoned Action, the behavioural intention of a student like any other technology user is influenced by attitude and behaviour [10]. Therefore, a student will behave (use the internet) based on the attitude and the intentions on the activity to be conducted using the internet technology. The intention to use a technology is also dependent on its perceived usefulness. According to TAM the intention to use a technology is based on the perceived usefulness of the technology and the perceived ease of use in the technology. On the other hand, the perceived usefulness of the technology defines the degree to which the technology will enable its users to enrich their performance at work and hence influence their attitude and behaviour towards the technology. The degree of acceptance of the technology is also seen to influence the context in which it will be used [4]. For students, the context in which internet technology is found, significantly influences their attitude and actions.

Student performance has been analysed using many methods. Some of these methods include Data mining techniques [11], Decision tree technique [12] and Factorization Techniques [13] among others. These approaches presume existence of databases with student attributes that influence performance. A new domain of knowledge in predicting student performance is developing around behavioural pattern analysis. This is considered to be richer in cognitive aspects of the students that largely influence the learning outcome of a student. Behavioural patterns have also been considered in the analysis of the student performance by [14]. [9] in a study to analyse student performance on a MOOCs platform, the data used was described by four V's (velocity, volume, variety and veracity). By considering these characteristics of data, they were pointing to the trend of moving away from static databases to more dynamic databases that handle student data. Moreover, this implies existence of unstructured databases whose data is generally non-linear in nature. In this paper, we consider the non-linear approach by modelling student behavioural patterns as a stochastic process with a number of random disturbances.

A. Stochastic Differential Equation (SDE) Student Performance predictive model

Student performance in a higher learning institution is influenced over time t≥0 by a number of factors which can be grouped as endogenous and exogenous factors. In this paper, we consider investment costs (I(t)) over a period of time t as exogenous factors. Student effort (E(t)) in the utilization of internet resources over the time period t and effectiveness (k) of the student effort over the time period t on the utilization of the internet resources to be the endogenous factors.

1. Investment:

Investment in the context of this paper is framed around the concept of behavioural costs. Behavioural costs captures two important elements that can be used to measure individual student investments in a learning process. These elements include: behavioural resources (also referred to as sacrifice) and opportunity costs. For a student to behave (act) in a particular way, he/she must make use of behavioural resources [15]. It is also important to note that for every outcome there is a comparable alternative outcome which the student must make a decision on based on the expected relative costs. [15] have identified these costs to be time costs (time budget), psychic costs (mental budget) and physical costs (physical budget).

In this paper, we introduce a fourth cost which we identify as technology costs. This cost touches on aspects to do with technology proficiency or technical attributes of the technology itself inherent on the student environment and the institutional technology utilization policies.

We thence define the measure of investment as:

$$\boldsymbol{I(t) = \sum f(t_c, M_c, P_c, T_c)} \qquad (1)$$

where t_c= time costs as measured on time- demand by an activity, M_c= Psychic costs as measured on the perceived mental demands of the activity, P_c = physical costs as expressed by the physical budget needed for the activity and T_c= the technology costs as expressed by the technological needs/resources/policies for the activity to be done. All these costs are relative except the time costs [16].

In a learning environment, the behavioural costs tend to exhibit unique characteristics due to the social nature of the environment and interactions that take place. Due to the likelihood of individual students borrowing a resource from one another and even seeking assistance from another, the individual costs tend to be shared out. Moreover, some costs are generally shared, for instance the physical costs. To even out these costs, we consider the concept of consumption ratios in the computations of the investment costs for individual students.

Thus equation (4) can be written as:

$$\boldsymbol{I(t) = \sum f\left(\frac{t_c}{t_b}, \frac{M_c}{M_b}, \frac{P_c}{P_b}, \frac{T_c}{T_b}\right) = \frac{Action\ price}{Action\ budget}} \qquad (2)$$

where t_b = budgeted time costs to undertake an activity, M_b= Perceived mental demand/Psychic costs to completion of an activity, P_b= Budgeted resources to undertake an activity and T_b = the budgeted/planned technology costs as expressed by the technological needs/resources/policies completion of an activity.

It is worthy to note the action price is influenced by attractiveness of a set goal. Less important goals tend to attract less budget and vice-versa. Therefore, student will invest more of their resources on what they consider important and the action price will not be significant due to the commitment and motivation.

1) Effort:

[17] demonstrated the relationship between attitude and actions. They considered that in a single act there is a corresponding attitude and context. They proposed a model for predicting behaviour outcome that considered attitude, beliefs, expected consequences and subjective norms. Subjective norms are important in a strictly controlled environment but in loosely bound environment like internet mediated learning environment, this factor may not play a significant role. This means, behavioural outcome is dependent on attitudes, beliefs and expected consequences. We therefore borrow this analogy to proxy the behavioural outcomes as a consequence of the actions of students in the utilization of internet technology in the context of learning.

[15] notes that attitude about an activity heralds the level of performance of the activity itself expressed as a behavioural intention which is stochastic in nature. In this paper, we therefore present a model that measures the student efforts in the utilization of internet technology as measured by considering the behavioural intentions and the student actions.

So, let the student behavioural intention be denoted by $BI(t)$ and the action that follows the intention be $A(t)$, then we can relate the two as:

$$\boldsymbol{BI(t) = e_t(A(t))} \qquad \boldsymbol{t \geq 0} \qquad (3)$$

where e_t is the relative effort put by the student to carry out a specific action.

Equation (3) can be written as:

$$\boldsymbol{e_t = \frac{BI(t)}{A(t)}} \qquad (4)$$

with, $A(t) = \sum_{t=1}^{n}(b_t * r_t)$

where b_t denotes the student belief in carrying out the activity, r_t the expected outcome from the activity.

Therefore, equation (4) become:

$$\boldsymbol{e_t = \frac{BI(t)}{\sum_{t=1}^{n}(b_t * r_t)}} \qquad (5)$$

Equation (5) gives a relative measure of the effort used by the student to achieve a particular behavioural outcome considering the actions taken at time t. It is assumed that the student behavioural intention will shift depending on a number of other controlling factors.

2) Effectiveness:

In this study effectiveness will measure the difference between two possible action strategies by the student. More often, students will use preferences in selecting the action. Effectiveness in this study is measured by:

$$\boldsymbol{k = A_1 - A_2 = w_1(EO_1 - EO_2) - w_2(BC_1 - BC_2)} \qquad (6)$$

where $A_1 - A_2$ is the relative student preference of action 1 to action 2 which measures the effectiveness of the actions selected. EO_1 , EO_2 represent the expected outcome from action 1 and 2 respectively. BC_1, BC_2 represent the behavioural costs associated with action 1 and 2 respectively.

3) Performance:

Assuming performance of student to be a continuous stochastic process, it can therefore be represented as a nonlinear stochastic system given by:

$$\boldsymbol{d(p(t) = f((p(t), i(t))dt + \sigma(t)d\omega(t),\ p(0) = p_0} \qquad (7)$$

where the performance state $p\epsilon\mathbb{R}^m$, the input $i\epsilon\mathbb{R}^m$, therefore drift term $f : \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}^m$ and the diffusion $\sigma: \mathbb{R}^n \rightarrow \mathbb{R}^m$.

The Weiner process ω in the system equation will model the randomness of the student performance due to unknown errors.

Now, let the student performance outcome be denoted by;

$$\boldsymbol{Y(t) = f(I(t), k, e(t))} \qquad \boldsymbol{Y_t \geq 0} \qquad (8)$$

Taking student performance over time t to be denoted by P_t; using the production function of the Cobb-Douglas type we model student performance outcome in (8) as:

$$\boldsymbol{P_t(t, Y(t)) = I_t^a(k_t e_t)^b} \qquad (9)$$

For some arbitrary but fixed student investment costs, $a\epsilon(0,1)$ and $b = 1 - a$.

The student effort k_t on the utilization of the internet technology is characterised by random dynamics due to the levels of competences on the internet technology and total productivity factor. It can therefore change over time following a geometric form of the Brownian motion represented by the stochastic differential equation (SDE) as:

$$\boldsymbol{dk_t = \mu_k k_t dt + \sigma_k k_t dw_t^k}, \quad \boldsymbol{k_0 > 0} \qquad (10)$$

Given the average improvement in performance $\mu_k \geq 0$ and constant unpredictability change in technological innovation, > 0 . $\mu_k k_t dt$ term represents the drift part and $\sigma_k k_t dw_t^k$represents the diffusion term with dw_t^k denoting the a Weiner process which has a Gaussian distribution characteristics.

The measures of the student effort can also be captured as geometric Brownian motion given by:

$$\boldsymbol{de_t = \mu_e e_t dt + \sigma_e e_t dw_t^e}, \quad \boldsymbol{e_0 > 0} \qquad (11)$$

For an average rate of change $\mu_e\epsilon\mathbb{R}$ and constant unpredictability of student effort $\sigma_e > 0$.

In this study, we consider that the investment level in technology by a higher learning institution is directly affected by the student demand for the technology and the perceived student performance due to the utilization of the technology. We further consider that the demand for internet technology in learning environment can be inferred by the effort put by the student on its usage. We therefore consider that universities are more likely to make policies on investment on the internet technology considering the effort put by the students to use the internet and expected performance. These dynamics are captured as:

$$\boldsymbol{dI_t = [P_t - \rho I_t - e_t c_t]dt + \sigma_I I_t dw_t^I}, \boldsymbol{I_0 > 0} \qquad (12)$$

where σ_I denotes the constant unpredictability investment, c_t is the random institutional utilization policies rate on internet technology (e.g. broadband sharing ratios, domain separation, platform accessibilities etc.) at time t and $\sigma_I > 0$.

Therefore, without loss of generality, c_t is dependent on $(I_m)_m \geq 0, (k_m)_m \geq 0, (e_m)_m \geq 0$ only and follows a Markovian property of memoryless time-homogeneous.

Hence,

$$\boldsymbol{c_t = c(I_t, k_t, e_t)} \; \forall \; \boldsymbol{institutional\ utilization\ policies}$$

Thus, equation (12) can be written as:

$$\boldsymbol{dI_t = [P_t - \rho I_t - e_t c(Y_t)]dt + \sigma_I I_t dw_t^I}, \boldsymbol{I_0 > 0} \qquad (13)$$

From the above equations, the values of I_t, k_t, e_t are all influenced by inherent random errors captured as uncertainties which are modelled as independent standard Brownian motions dw_t^I, dw_t^k and dw_t^e respectively.

Equation (13) can be rewritten as:

$$\boldsymbol{dI_t = \left[I_t^a(k_t e_t)^b - \rho I_t - e_t c(Y_t)\right]dt + \sigma_I I_t dw_t^I}, \quad \boldsymbol{I_0 > 0} \qquad (14)$$

Assuming a student as a consumer of the internet technology is supposed to have a constant rate $\varphi \geq 0$ of the time-preference and Constant Relative Risk Aversion (CRRA) utility function given by

$$\boldsymbol{u(c_t) = \frac{c_t^{1-\theta}-1}{1-\theta}} \qquad (15)$$

By considering work by [18], we assume $\theta = a$ under circumstances where we take uniformity in the applicability of intervening factors like resources accessibility policies.

In this study, we assume internet technology utilization by students in a higher learning institutions is dependent on a collection of institutional policies and internet consumption strategies which we denote by $C(t, y)$ with t representing time point and y is the observable value of Y_t.

The desire of many learning institutions is to optimize consumption rate of the internet technology as a resource by its student. They thus set utilization strategies and policies as estimated by the utility function given by equation (15) aimed at aiding students to achieve best performance outcomes. A study by [19] estimates the optimal consumption utility value by:

$$\boldsymbol{c(Y_t) = h\frac{I_t}{e_t} \qquad g \geq 0} \qquad (16)$$

And by the principle of marginal rate of technical substitution (MRTS) $h = \frac{b}{a}$. It can be noticed that the optimal consumption rate of the internet technology is dependent on investment costs and the effort.

Substituting equation (16) in equation (14) gives:

$$\boldsymbol{dI_t = \left[I_t^a(k_t e_t)^b - \rho I_t - hI_t\right]dt + \sigma_I I_t dw_t^I, \quad I_0 > 0} \qquad (17)$$

Which simplifies as:

$$\boldsymbol{dI_t = \left[I_t^a(k_t e_t)^b - (\rho + h)I_t\right]dt + \sigma_I I_t dw_t^I, \quad I_0 > 0} \qquad (18)$$

Considering equation (9) and applying Ito's formula, we get:

$$\boldsymbol{dP_t = d\left(I_t^a(k_t e_t)^b\right)}$$
$$\boldsymbol{= (k_t e_t)^b dI_t^a + I_t^a d(k_t e_t)^b + dI_t^a d(k_t e_t)^b}$$
$$\boldsymbol{= f\left(t, (I_t^a k_t e_t)^b\right)dt + (k_t e_t)^b I_t^a \sigma(t)^{tr} G(t) dw_t}$$
$$\boldsymbol{= \left(aI_t^{a-1}(k_t e_t)^b + b\, I_t^a\,(k_t e_t)^{b-1}\right)dt + I_t^a(k_t e_t)^b \sigma(t)^{tr} G(t) dw_t} \qquad (19)$$

Which is the stochastic differential equation (SDE) with initial condition $P_0 = x$. Equation (19) gives the performance of a student at time t.

A. SDE student performance Predictive model parameters

We consider the dynamics surrounding the student learning environment to be of continuous nature and therefore the models used are continuous-time stochastic differential equations. Therefore, a Maximum Likelihood Estimation (MLE) method was used to estimate the parameters. **Table 1** gives parameters that have been used recently in some growth studies as presented by [19]. We adopted these parameters into the model developed.

Table 1. SDE Student Performance Predictor Model Parameters

Parameter	Value	Reference
a	0.1-0.77	[20]
σ_k	0.0148	[21]
σ_I	0.12	[22]
σ_e	0.01	[23]
μ_E	0.0176	[24]
μ_e	0.01-0.02	[23]
ρ	0.05-0.08	[23]
θ	1.0-10.0	[23]

I. Results and Discussion

To test on the performance of the model, an observational study was conducted on a university student over a limited period of time. The aspects that were observed included assignment/homework taken using internet mediated platform (Moodle), time spent on the internet, resources accessed, time budgeting for online research, student skill set on Internet Technology(self-efficacy), the attitudes of the student towards Internet Technology, the nature and availability of the Internet Technology both at school and home, environment of internet usage, availability of collaborations on the internet mediated platforms, institutional internet usage guidelines and student internet technology preferences.

Analysis of the data was done using a Continuous Time Stochastic Modelling in R (CTSMR) version 3.4.2 which has the capability of handling non-linear stochastic processes.

Table 2 shows the estimated parameter values using the MLE technique.

Table 2. Estimated model parameters

Parameter	True Value	Estimated Value
a	0.1-0.77	0.76905
σ_k	0.0148	0.0148
σ_I	0.12	0.012
σ_e	0.01	0.01
μ_E	0.0176	0.0176
μ_e	0.01-0.02	0.014142
ρ	0.05-0.08	0.063246

Figure 1 shows the predicted student performance trajectories. It was observed that with near similar investment, student performance improved over time as student effort reduced. This could be attributable to the fact that as the student continued using the internet for learning, there was marked increase in proficiency and ability to utilize the Internet Technology to perform similar tasks.

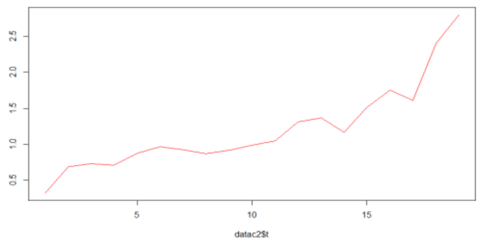

Figure 1. Student Performance Trajectory

III. CONCLUSION

Predicting student performance is important to the learners and education practitioners. This will help them to understand the implications of whatever resources, strategies and policies they employ. This paper has analysed the implications of the behavioural patterns of student by focusing on student effort, investment costs and effectives of the strategies they employ in achieving good performance. The paper presents an SDE student performance predictive model which considers student performance as a stochastic process characterised by random noises. It therefore provides educational practitioners with a framework of looking at performance as process rather than an event.

IV. REFERENCES

[1]. Nath, K, Dhar, S, & Basishtha, S. (2014). Web 1.0 to Web 3.0—evolution of the Web and its various challenges. International Conference on Reliability, Optimization and Information Technology (ICROIT) (pp. 86–89). Faridabad: IEEE. doi:10.1109/ICROIT.2014.6798297.

[2]. Kop, R. (2011). The challenges to connectivist learning on open online networks: learning experiences during a massive open online course. In The International Review of Research in Open and Distance Learning, Special Issue - Connectivism: Design and Delivery of Social Networked Learning, 12(3).

[3]. Bradley, J, Barbier, J, & Handler, D. (2013). Embracing the Internet of Everything to capture your share of $ 14.4 trillion.

[4]. Omwenga V. (2016). Context-Centered Mobile Application Development for Effective Adoption of Mobile Technology. Journal of Computer Engineering Vol. 18, Issue 1, 01-07.

[5]. Usamah M., Norlinda B. Pauziah M., and Rosni K. (2013). An Ovreview of Using Academic Anlytics to Predict and Improve Students' Achievement: A proposed Proactive Intelligent

Intervention. IEEE 5th Conference on Engineering Education (ICEED)

[6]. Spennemann, D. H.R. (2007). Learning and teaching 24/7: daily internet usage patterns at nine Australian universities. Campus-Wide Information Systems, Vol. 24 Issue: 1, pp.27-44, doi: 10.1108/10650740710726473

[7]. Anunobi, C., & Ogbonna, A. (2013). Web 2.0 use by librarians in a state in Nigeria. Developing Country Studies, 2(5), 7-66. Osunade, O., & Ojo, O. (2012). Library and internet usage: a case study of the University of Ibadan. The Information Technologist, 3(2), 19-29.

[8]. Hughes G. and Dobbins C. (2015). The utilization of Data Analysis Techniques in Prediction Student Performance in Massive Open Online Courses (MOOCs). Research and Practice in Technology Enhanced Learning 10:10

[9]. Ajzen, I. (1991). The theory of planned behavior. Organ. Behav. Hum. Decis. Process. 50 (2), 179–211

[10]. Amirah M. S., Whidah H. and Nur'aini A. R. (2015). A Review on Predicting Student's Performance using Data Mining Techniques. Procedia Computer Science 72, 414-422

[11]. Anjali B. R. and Ankita A. N. (2017). Student Performance Prediction Using Decision Tree Technique. International Journal of Computational Intelligence Research 13 (7) 1735-1741

[12]. Nguyen T. Lucas D., Tom´aˇs, H., Artus K., Alexandros N., and Lars S. (2011). Factorization Technique for Predicting Student Performance. Educational Recommender Systems and Technologies: Practices and Challenges. 10.4018/978-1-61350-489-5.ch006.

[13]. Anderson, A, Huttenlocher, D, Kleinberg, J, & Leskovec, J. (2014). Engaging with massive online courses. In Proceedings of the 23rd International Conference on World Wide Web (WWW '14) (pp. 687–698). New York: ACM Press. doi:10.1145/2566486.2568042.

[14]. Verhallen T. M. and Pieters R. (1984). Attitude theory and Behavioural Costs. Journal of Economic Psychology 5(3), 223-249

[15]. Shugan S.M. (1980). The cost of thinking. Journal of Consumer Research 7, 99-111

[16]. Ajzen I. and Fishbein M. (1977). Attitude-behaviour: A theoretical analysis and review of empirical research. Psychological Bulletin 84, 888-918.

[17]. Posh O. and Wälde K. (2010). On the Non-Causal Link Between Volatility and Growth. CREATES Economics Working Paper 2009-10, University of Aarhus, School of Economics and Management

[18]. Feicht R. and Stummer W. (2010). Complete Closed-form Solution to Stochastic Growth Model and Corresponding Speed of Economic Recovery. IWQW Discussion Paper Series, No. 05/2010

[19]. Acemoglu D. and Guerrieri V. (2008). Capital Deepening and Non-Balanced Economic Growth. Journal of Political Economy, 116(3): 467-498

[20]. Bucci A., Colapinto C., Forster M. and Torre D. L. (2008). On Human Capital and Economic Growth with Random Technology Shocks (To check further on the journal)

[21]. Palacios, M. (2010). Human capital as an asset class: Implications from a general equilibrium model. Working paper, Vanderbilt University, Owen Graduate School of Management

[22]. Jensen B. S. and Richter M. (2007). Stochastic One-sector and Two-sector Growth Models in Continuous Time. Stochastic Economic Dynamics, Part 5 pp 167-216. Business School Press Copenhagen

[23]. Williams N. (2004). Small Noise Asymptotics for a Stochastic Growth Model. Journal of Economic Theory, 119(2): 271-298

[24]. King R. G. and Robelo S. T. (1993). Transitional Dynamics and Economic Growth in the Neoclassical Model. American Economics Review, 83(4): 908-31