Volume LX No. 2
December 2023

p-ISSN 1655-1516
e-ISSN 2984-8156

The Philippine Review of Economics

ARTICLES IN THIS ISSUE

A joint publication of the
University of the Philippines
School of Economics
and the **Philippine Economic Society**

The Philippine Review of Economics

A joint publication of the UP School of Economics (UPSE)
and the Philippine Economic Society (PES)

Aims and Scope: The *Philippine Review of Economics* (PRE) invites theoretical and empirical articles on economics and economic development. Papers on the Philippines, Asian and other developing economies are especially welcome. Book reviews will also be considered.

The PRE is published jointly by the UP School of Economics and the Philippine Economic Society. Its contents are indexed in the *Journal of Economic Literature*, EconLit, and RePec. PRE's readership includes economists and other social scientists in academe, business, government, and development research institutions.

Publication Information: The PRE (p-ISSN 1655-1516; e-ISSN 2984-8156) is a peer-reviewed journal published every June and December of each year. A searchable database of published articles and their abstracts is available at the PRE website (http://pre.econ.upd.edu.ph).

Subscription Information:
Subscription correspondence may be sent to the following addresses:

- css@pssc.org.ph
- PSSC Central Subscription Service,
  PSSCenter, Commonwealth Avenue, 1101, Diliman,
  Quezon City, Philippines.
  P.O. Box 205, UP Post Office, Diliman, Quezon City,
  Philippines 1101
  PHONE: 922-9627, FAX: 924-4178/926-5179

Submissions: Authors may submit their manuscripts to the addresses below:

- pre.upd@up.edu.ph
- The Editor, The Philippine Review of Economics, Rm 237, School of Economics, University of the Philippines, Diliman, Quezon City, 1101.

Manuscripts must be written in English and in MS Word format. All graphs and tables must be in Excel format. Submission of a manuscript shall be understood by the PRE as indicating that the manuscript is not under consideration for publication in other journals. All submissions must include the title of the paper, author information, an abstract of no more than 150 words, and a list of 3–4 keywords. Complete guidelines can be viewed in the PRE's website.

Acknowledgements: The PRE gratefully acknowledges the financial support towards its publication provided by the Philippine Center for Economic Development (PCED). The Review nonetheless follows an independent editorial policy. The articles published reflect solely the editorial judgement of the editors and the views of their respective authors.

The Philippine Review of Economics

Vol. LX No. 2
December 2023

p-ISSN 1655-1516
e-ISSN 2984-8156
DOI: 10.37907/ERP3202D

The Philippine Review of Economics
60(2):1-24. DOI:10.37907/1ERP3202D

Shared prosperity characterized by four development goals: pro-poor growth, pro-poor development, inclusive growth, and inclusive development*

Nanak Kakwani**
University of New South Wales

Zakaria Siddiqui
Jamia Millia Islamia University

This paper is on shared prosperity and its measurement. Economic growth enhances total prosperity, increasing the economic pie in society, but the pie distribution determines how the population shares the pie. Based on a social welfare framework, we have developed an integrated methodology to evaluate growth and distribution simultaneously. Linking the two phenomena gives rise to four development goals: (i) pro-poor growth, (ii) inclusive growth, (iii) pro-poor development, and (iv) inclusive development. These four goals provide an alternative characterization of shared prosperity. The paper defines the four goals, providing a methodology to operationalize them using real-world data. The empirically measured goals inform at what rate the shared prosperity is enhancing in any country or the world. The methodology is applied globally to determine whether the growth and development have been pro-poor and inclusive in 173 countries over the two decades in the new millennium.

JEL classification: D63, D31, O11, O20, O47
Keywords: shared prosperity, pro-poor growth, inclusive growth and development, poverty, inequality

1. Introduction

In the 1950s and 1960s, trickle-down was the dominant development strategy for bettering people's lives. It implied that economic growth was the dominant factor that would automatically enhance people's living standards. The growth process, resulting from market forces, generally benefits the wealthy first, and then in the second round, the poor benefit when the rich start spending their gains from growth. The trickle-down ensures a vertical flow of the benefits of growth

* Lecture delivered at the Philippine Economic Society 61st Annual Meeting on November 8, 2023.
** Address all correspondence to n.kakwani@unsw.edu.au.

from the rich to the poor. Thus, economic growth only benefits the poor indirectly through vertical flows from the rich. The trickle-down was silent on how much benefits of growth flow to the poor. The rich may reap huge benefits, but at the same time, the poor may receive only a meager fraction of the total benefits.

Thus, the view in development economics was that the government's strategy should promote investments, increase production capabilities, and enhance economic growth. The governments need not be concerned with how economic growth distributes benefits among the people, and the distribution was not considered a fundamental problem for serious study.

In his book, Bronfenbrenner [1971] raised an important question, "Is distribution a sufficiently important problem for serious study, and if so, why?" Chapter 1 of his book presents a representative sample of divergent views of economists. Some economists viewed distribution as fundamental, while others thought that distribution was unimportant. There existed little consensus among them.

Economic growth provides means, but distribution is fundamental to economic and social equality. In this context, the following quotation from Sen and Drèze [1989] is helpful: "*Economic growth is very important as a means for bettering people's lives, but to go much faster, it has to be combined with devoting resources to remove illiteracy, ill health, undernutrition, and other deprivations.*"

America has been the wealthiest economy in the world; recently, the Nobel Laureate economist Angus Deaton (June 7, 2023) has emphasized the flipside of American progress, calling it economic failure or failed economics. He argues, "growth is worthless to those who do not share it. [Gross domestic product or] GDP is blind to who benefits and who loses, and over the last half-century, most Americans have not seen the growth in incomes that might seem warranted by the growth in the economy." Thus, Deaton has forcefully argued that we cannot achieve prosperity for all through economic growth without considering the distribution of the output generated by economic growth.

This paper links the two phenomena of growth and distribution that give rise to four development goals: (i) pro-poor growth, (ii) inclusive growth, (iii) pro-poor development, and (iv) inclusive development. The literature has not distinguished the four development goals. We view these four goals as the alternative characterizations of shared prosperity.

This paper's main contribution is defining the four goals and providing a social welfare framework to operationalize them using real-world data. Thus, these goals offer alternative ways of measuring shared prosperity.

Based on a social welfare framework, we have developed an integrated methodology to evaluate the size of the pie and its distribution simultaneously. Our proposed social welfare framework links economic growth and distribution, giving rise to the four development goals through this linkage. From this framework, we can also determine the contributions of growth and distribution to social welfare and well-being. This decomposition is essential to understanding the policy implications of shared prosperity.

The empirically measured goals inform at what rate the shared prosperity is enhancing in any country or the world. Thus, the paper's main contribution is to provide an operational system to monitor shared prosperity.

The methodology developed in the paper is applied globally to determine whether the growth and development have been pro-poor and inclusive in 173 countries over the two decades in the new millennium.

2. What is pro-poor growth?

The term pro-poor growth is relatively new and emerged in the late 1990s. Many development practitioners began discussing it but did not offer a precise concept of pro-poor growth. International agencies such as the UN [2000] and the OECD [2001] defined pro-poor growth as growth that benefits the poor and provides them with opportunities to improve their economic situation. *The poverty reduction strategy* by the Asian Development Bank describes pro-poor growth as labor-absorbing growth accompanied by policies and programs that mitigate inequalities and facilitate income and employment generation for the poor, particularly women and other traditionally excluded groups. These definitions are very broad and focused on policies to achieve pro-poor growth. Before discussing policies, it makes logical sense to define pro-poor growth precisely. The broad policies are not a helpful guide in measuring pro-poor growth.

Kakwani and Pernia [2000] developed the precise concept of pro-poor growth arguing that pro-poor growth is biased in favor of the poor, meaning that the poor must enjoy higher benefits of growth than the non-poor. Based on this definition, they proposed an operational measure of pro-poor growth, which informed when one could say that growth is pro-poor. And if so, to what degree?

Kakwani and Son [2008] proposed two alternative definitions of pro-poor growth. A brief review of these definitions is now provided.

(i) Relative definition of pro-poor growth

The growth is relative pro-poor (anti-poor) if the average relative growth rate of the income is positive, and the poor benefit proportionally more (less) than the non-poor.

The growth is also relative pro-poor (anti-poor) if the relative growth rate of income is negative, and the poor suffer a proportionally smaller (larger) decline in their income than the non-poor.

Kakwani and Pernia [2000] proposed this definition, implying that growth results in income redistribution favoring the poor. This is a relative concept of pro-poor growth because the growth process reduces relative inequality.

(ii) Absolute definition of pro-poor growth

The growth is absolute pro-poor (anti-poor) if the average absolute growth of income is positive, and the poor benefit absolutely more (less) than the non-poor.

The growth is also absolute pro-poor (anti-poor) if the average absolute growth rate of income is negative, and the poor absolutely suffer a smaller (larger) decline in their income than the non-poor.

Kakwani and Son [2008] proposed this definition, implying that growth results in the redistribution of income in favor of the poor, contributing to a greater absolute gain of income for the poor than the non-poor. If the growth is negative, the redistribution of income due to growth leads to a smaller loss of absolute income for the poor than for the non-poor. This is an absolute concept of pro-poor growth because the growth process reduces the absolute inequality of income. Kolm [1976] developed the idea of absolute inequality, which remains unchanged when everyone's income changes by the same amount. This paper has extended this idea to measuring absolute pro-poor growth.

3. Poverty equivalent growth rate (PEGR) explained

The linkage between growth and poverty is complex and is determined by inequality changes. Thus, pro-poor growth provides the interrelationship between three factors: poverty, inequality, and growth, known in the literature as the PIG axis [Sumner 2003]. Kakwani and Son [2008] developed the idea of a "poverty equivalent growth rate" (PEGR) that takes into account both the growth rate in mean incomes and how the benefits of growth are distributed among the poor and non-poor. It encompasses the two definitions of pro-poor growth discussed in the previous section. This paper demonstrates that the PEGR satisfies an essential requirement that the magnitude of poverty reduction is a monotonically increasing function of the PEGR. Thus, the PEGR is an effective tool to reduce or alleviate poverty; maximization of the PEGR implies a maximum reduction in poverty. The government's social objective should be to maximize the PEGR.

The derivation of the PEGR is explained in [Kakwani and Son 2008]. The following hypothetical example can provide an intuitive explanation of the PEGR. Suppose the actual growth rate is seven percent, which has reduced poverty by ten percent, meaning that $\delta = -0.10$ and $\gamma = 0.07$. Suppose the growth elasticity of poverty is $\eta = -1.2$, interpreted as a one percent increase in mean income reduces poverty by 1.2 percent, provided the relative inequality had not changed. The growth in poverty under the counterfactual that inequality had not changed would be $-1.2 \times 0.07 = -0.084 \approx -8.4$ percent. The actual poverty reduction is ten percent, meaning that the actual poverty reduction is higher than the reduction that would have occurred if growth were inequality neutral, which gives a pro-poor index $\varphi = (-.10)/(-.084) = 1.19$. Hence, the poor enjoy 19 percent higher benefits than the non-poor, so growth is pro-poor. The PEGR $= 0.07 \times 1.19 = 0.08 \approx 8$ percent, which

is higher than the actual economic growth rate of seven percent. Thus, there is a gain of one percent in the growth rate because growth is pro-poor.

Suppose the economy suffered a recession, so the economic growth rate declined by five percent, implying $\gamma = -0.05$, which led to an increase in poverty by seven percent, giving $\delta = 0.07$. If the recession were inequality neutral, poverty would have increased by $-1.2 \times (-0.05) = 0.06 \approx 6$ percent. The actual increase in poverty is seven percent, which yields the pro-poor index $\varphi = 7/6 = 1.17$. It means that the poor suffer a 17 percent higher loss of income than the non-poor; therefore, the recession is anti-poor. Thus, the PEGR $= -0.05 \times 1.17 = -0.059 \approx -5.9$ percent, which is lower than the actual growth rate of -5 percent. Therefore, society suffers a loss of growth rate equal to 0.9 percent. A similar interpretation applies to the absolute PEGR.

This hypothetical example has a critical message: it shows that pro-poor growth contributes to a gain in the growth rate in poverty reduction, while anti-poor growth results in the loss of the growth rate in poverty reduction. This result is intuitive and can be more readily conveyed to policymakers.

4. Poverty social welfare approach to pro-poor growth

The PEGR requires the specification of the poverty line and an aggregate poverty measure, and several poverty measures are available in the literature based on alternative assumptions. The PEGR can be calculated for any poverty measure, a general method encompassing any poverty measure. Any income and expenditure household survey can be used to operationalize the technique. This technique requires the estimation of the growth elasticity of poverty η, and Kakwani and Son [2008] proposed to estimate the elasticity using the poverty decomposition proposed by Kakwani [2000]. Many researchers have found the estimation of this elasticity rather tricky. This section offers an alternative method of estimating pro-poor growth using the poverty social welfare approach. The poverty social welfare function can be directly linked to poverty measures, implying a one-to-one relationship. But, at the same time, it provides a more straightforward way of identifying a growth pattern as pro-poor.

The idea of poverty social welfare function is explained below.

Suppose z is the poverty line, the income below which individuals cannot satisfy their minimum needs. Persons are identified as poor if their income x is below the poverty line. We develop below a general class of poverty social welfare functions and show how it can drive a class of pro-poor growth indices.

Suppose $v_k(z, x)$ is the weight given to a poor person with income x, when poverty line is z, defined as

$$\begin{aligned} v_k(z,x) &= \frac{(k+1)}{H}\left[\frac{H-F(x)}{H}\right]^k \quad if\ x < z \\ &= 0 \qquad\qquad\qquad\qquad\quad if\ x \geq z. \end{aligned} \tag{1}$$

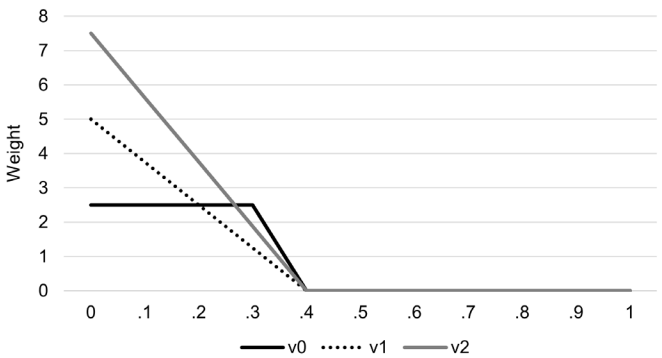
$F(x)$ is the probability distribution function, which is the probability of a person with income less than x. H is the proportion of poor identified by the poverty line z.

The total weight in the domain of x adds up to one:

$$\int_0^H v(z,x)\,f(x)dx = \frac{(k+1)}{H}\int_0^H \left[\frac{H-F(x)}{H}\right]^k f(x)dx = 1. \qquad (2)$$

The poorest person gets the maximum weight of $(k + 1)$, which decreases monotonically as income increases and becomes zero when the income of the poor increases to the poverty line z or higher. Thus, all the weight is given to the poor, and the non-poor receives zero weight, which characterizes poverty social welfare functions. Figure 1 depicts the weights assigned to the poor. For an illustration, H is assumed to be 0.4.

FIGURE 1. Poverty social welfare functions



Note: Weights: H = 0.4

The figure depicts the three alternative weighting schemes. When $k = 0$, every poor person receives the exact weight of 2.5 until the income of the poor equals the poverty line so that all the non-poor receive zero weights. When $k = 1$ or $k = 2$, the weight decreases monotonically as the income of the poor increases, attaining the value 0 when the poor cross the poverty line. This weighting scheme leads to the following class of poverty social welfare functions.

$$x^*(z,k) = \frac{1}{H}\int_0^z xv(z,x)\,f(x)dx = \frac{(k+1)}{H}\int_0^z x\left[\frac{H-F(x)}{H}\right]^k f(x)dx\,, \qquad (3)$$

which is the money metric poverty social welfare function measured in the income currency such as the dollar. This social welfare class depends on the income ranking of the poor. Sen [1976] proposed the idea of rank order ranking from the viewpoint of capturing the relative deprivation suffered by persons when they compare their economic circumstances with others in society. The basic intuition behind the rank ordering is that the lower a person is on the welfare scale, the higher this person's

sense of deprivation. Intuitively, the person suffering from the highest deprivation must receive the most importance and, thus, the largest weight.

When $k = 0$, $x^*(z, k)$ becomes

$$x^*(z, 0) = \frac{1}{H}\int_0^z xf(x)dx\,, \tag{4}$$

which equals the mean income of the poor. It is the most straightforward poverty social welfare function. This social welfare function has one limitation: it gives equal weight to all people experiencing poverty, irrespective of economic circumstances. All poor cannot be identical; they have different incomes, so they must have different weights. Figure 1 shows that when $k > 0$, the importance given to the poor decreases linearly as their income increases. As k increases from one to two, Figure 1 also shows that the weight function becomes steeper, giving relatively greater weight to the poorer persons among the poor. It means that the parameter k is interpreted as the inequality aversion parameter; as k rises, more and more importance is given to transfers among the poor at the lower end of the distribution and less weight to the transfer at the top. This is a desirable property if society is concerned with giving greater importance to poorer persons among the poor. Thus, it would be more appropriate to measure pro-poor growth using the general class of poverty social welfare functions in (3) for $k > 0$; the higher the value of k, the greater society's inequality aversion.

Having explained the poverty social welfare function, we show how we can derive the measures of pro-poor growth from it.

Suppose $\gamma = \Delta\ln(\mu)$ is the relative growth rate of the mean income of the society, which can be shown to give equal proportion weight to everyone in society. Further, suppose $\gamma(z, k) = \Delta\ln(x^*(z, k))$ is the growth rate of the social welfare $x^*(z, k)$, which gives all the weight to only the poor, with the poorest getting the maximum weight. We may now define the pro-poor index as follows.

If $\gamma(z, k) > \gamma$, the growth will be pro-poor because the growth will benefit the poor proportionally more than the non-poor [definition (i) of pro-poor growth]. That leads to a relative pro-poor index $\rho(z, k)$ given by

$$\rho(z, k) = \frac{\Delta\ln(x^*(z, k))}{\Delta ln(\mu)} = \frac{\gamma(z, k)}{\gamma}\,, \tag{5}$$

where $\gamma(z, k)$ is the relative growth rate of the poverty social welfare $x^*(z, k)$. Poverty social welfare function gives the highest weight to the poorest person in society, and the weight decreases monotonically with income, becoming zero as the person's income becomes equal to or higher than the poverty line z. The non-poor persons get zero weight, which implies that the growth will be pro-poor (anti-poor) if the growth in social welfare $\gamma(z, k)$ is higher (lower) than the growth in the mean of society γ, because the poor will receive greater (smaller) proportional growth benefits.

Suppose $\gamma > 0$; growth will be pro-poor (anti-poor) if $\rho(z, k)$ is greater (smaller) than unity. If $\gamma < 0$, the growth will be pro-poor (anti-poor) if $\rho(z, k)$ is smaller (greater) than one because people experiencing poverty suffer a smaller (larger) loss of income due to the downturn in the economy.

The pattern of relative growth is determined by

$$\gamma(z, k) = \gamma + (\rho(z, k) - 1)\gamma, \tag{6}$$

which immediately shows that there will always be a gain (loss) in the relative growth of poverty social welfare if the growth process is pro-poor (anti-poor). The decision rule regarding the gain or loss in growth rate is straightforward to explain to the policy makers: the gain signifies pro-poor growth, and the loss anti-poor growth.

Similar to the relative pro-poor index in (3), we can also define an absolute pro-poor index for the class of social welfare function, $x^*(z, k)$ as

$$\rho^*(z, k) = \frac{\Delta x^*(z, k)}{\Delta \mu} = \frac{\gamma_A^*(z, k)}{\gamma_A}. \tag{7}$$

From definition (ii), the growth is absolute pro-poor (anti-poor) with an absolute positive growth rate, γ_A; the poor receive greater (smaller) absolute benefits than the non-poor, implying that $\rho^*(z, k)$ is greater (smaller) than one. Similarly, if $\gamma_A < 0$, the growth is absolute pro-poor (anti-poor) if the absolute loss of growth for the poor is smaller (larger) than that of the non-poor, implying that $\rho^*(z, k) < 1$ [$(\rho^*(k) > 1$], respectively.

The pattern of absolute growth is determined by

$$\gamma_A^*(z, k) = \gamma_A + (\rho^*(z, k) - 1)\gamma_A, \tag{8}$$

which immediately shows that there will always be a gain (loss) in the absolute growth of social welfare if the growth process is absolute pro-poor (anti-poor).

5. What is pro-poor development?

First, we need to clarify what development is. It is a complex issue, having different meanings for different people, and economic growth is commonly perceived as development. If a country achieves high economic growth, it is applauded as a country with a high level of development. Economic growth is measured in income space, which provides people with the means to lead a better life. Means are necessary but insufficient to give people the quality of life they must have.

According to Nobel Laureate Amartya Sen [1983], economic development has to be concerned with the kind of life people can lead and what they can or cannot do;

for example, whether they are well-nourished, get an education, or able to escape avoidable morbidity. His idea of development relates to enhancing people's well-being (or standard of living). He developed the most comprehensive framework of well-being through functionings and capabilities. While functioning is people's achievement, capability is their ability to achieve. Functionings are directly related to what life people lead, whereas capabilities are related to people's freedom in choosing the functions they value. Thus, development is a multidimensional concept defined in terms of capabilities that reflect the extent of freedom people have in determining the life they wish to lead. Following this framework, we describe development as enhancing peoples' capabilities (well-being).

Economic growth generates people's incomes, which are the means enabling people to have a command over commodities. But Sen's idea of well-being relates to the kind of life people can lead. Thus, well-being is the people's ultimate achievement, which we call ends, whereas means generated by economic growth enable people to achieve these ends. It is essential to note that there is no one-to-one relationship between means and ends.

We define development more narrowly as ends, whereas economic growth is a means. Means and ends have different characteristics; means can impact ends, so they are related, but still distinct, and policies to enhance means will differ from those that enhance ends.

This paper defines and measures four development goals: pro-poor and inclusive growth based on means and pro-poor and inclusive development on the ends. So, we treat them as different development goals.

Well-being is a multidimensional concept reflecting many aspects of people's lives. Several indicators measure well-being, and constructing a composite index to measure overall well-being is not essential. The construction of a composite well-being index suffers from many conceptual issues, well-documented in the literature [Kakwani and Son 2022]. Constructing a composite index requires weights to be assigned to various dimensions of well-being, and no meaningful method exists for determining the weights. We have refrained from creating composite indices of pro-poor development. Our conclusions on pro-poor development derive from the individual development indicators, which are sensible approaches to formulating policies.

Pro-poor development concerns the performance of the poor in achieving development relative to the non-poor. We propose the following two definitions of pro-poor development.

(iii) Relative pro-poor development

The development is relatively pro-poor (anti-poor) if the average relative well-being growth rate is positive and the poor enjoy a proportionally higher (lower) increase in well-being than the non-poor.

The development is also relatively pro-poor (anti-poor) if the average relative well-being growth rate is negative and the poor suffer a proportionally lower (higher) decline in well-being than the non-poor.

(iv) Absolute pro-poor development

The development is absolute pro-poor (anti-poor) if the average absolute well-being growth rate is positive, and the poor enjoy an absolutely higher (lower) increase in well-being than the non-poor.

The development is also absolute pro-poor (anti-poor) if the absolute well-being growth rate is negative and the poor suffer an absolutely lower (higher) decline in well-being than the non-poor.

6. The measurement of pro-poor development

Suppose $\omega(x)$ is the well-being indicator of a person with income x; several indicators characterize the overall well-being. For ease of presentation, $\omega(x)$ will be referred to as well-being.

We propose generalizing the poverty social welfare function in (3) to achieve this objective. This generalization will be called Poverty Social Well-being Function (PSWBF) given by

$$\omega_P^*(z, k) = \frac{(k+1)}{H}\int_0^z \omega(x)\left[\frac{H - F(x)}{H}\right]^k f(x)dx, \tag{9}$$

which links the well-being with the economic circumstances of the poor.

When $k = 0$, $\omega_P^*(k)$ collapses to $\bar{\omega}_z$ given by

$$\bar{\omega}_z = \frac{1}{H}\int_0^z \omega(x)f(x)dx, \tag{10}$$

which is the mean well-being of the poor. This is the most straightforward poverty social well-being function. Its main limitation is that the well-being of all the poor gets the same weight irrespective of their economic situation. However, if $k > 0$, the weight given to the well-being of the poor varies with their income. The well-being of the poorest gets the highest importance.

The pro-poor relative development index for the (PSWF) is given by

$$\tau_P(z, k) = \frac{\Delta Ln(\omega_P^*(z, k))}{\Delta Ln(\bar{\omega})} = \frac{\sigma_P(z, k)}{\sigma}, \tag{11}$$

where $\sigma_P(z, k)$ is the relative growth rate of poverty social well-being, and σ is the relative growth rate of the well-being of the whole population. The development, based on definition (iii), will be relative pro-poor (anti-poor) if $\sigma > 0$, *and* $\tau_P(z, k)$ is greater (less) than one. If $\sigma < 0$, *and* $\tau_P(z, k)$ is less (greater) than one, the development will be relatively pro-poor (anti-poor).

The pattern of pro-poor development is described by

$$\sigma_P(z, k) = \sigma + (\tau_P(z, k) - 1)\sigma, \tag{12}$$

which immediately shows that relative pro-poor development leads to a gain in relative well-being growth rate, while anti-poor development results in a loss in relative well-being growth rate. Thus, we propose to measure the degree of relative pro-poor development by the gain or loss of relative growth in a well-being indicator.

The pro-poor absolute development index for the PSWF is given by

$$\tau_P^*(z, k) = \frac{\Delta(\omega_P^*(z, k))}{\Delta(\bar{\omega})} = \frac{\sigma_P^*(z, k)}{\sigma^*}, \tag{13}$$

where $\sigma_P^*(z, k)$ is the absolute growth rate of poverty social well-being, and σ^* is the absolute growth rate of the well-being of the whole population. The development, based on definition (iv), will be absolute pro-poor (anti-poor) if $\sigma^* > 0$, *and* $\tau_P^*(z, k)$ is greater (less) than one. If $\sigma^* < 0$, *and* $\tau_P^*(z, k)$ is less (greater) than one, the development will be relatively pro-poor (anti-poor); and will be absolute pro-poor (anti-poor) if $\tau_P^*(k)$ is greater (less) than one. The pattern of pro-poor development is described by

$$\sigma_P^*(z, k) = \sigma^* + \tau_P^*(z, k) - 1)\sigma^*, \tag{14}$$

which immediately shows that absolute pro-poor development leads to a gain in absolute well-being growth rate, while anti-poor development results in a loss in well-being growth rate. Thus, we propose to measure the degree of absolute pro-poor development by the gain or loss of absolute growth of a development indicator.

7. What is inclusive growth, and how did it evolve?

What is the origin of the term inclusive growth? Our simple answer is that we do not know, and our Google search did not help. The development literature, however, has integrated the concept of inclusive growth into policymaking. In the new millennium, there has been widespread debate on the idea, still providing no clear definition of what inclusive growth is and how it differs from other development ideas proposed in the literature. The concept remains elusive, as pointed out by Ranieri and Ramos [2013] in a one-pager publication of the International Policy Centre for Inclusive Growth. A careful review of various ADB documents revealed many conflicting definitions of inclusive growth, as pointed out by Klassen [2010]. He concluded that inclusive growth concepts are vague and do not allow easy quantitative operationalization. Further complicating matters, the World Bank defines inclusive growth in ways that are at odds with the ADB concept.

The debate on inclusive growth advanced in India. Unfortunately, it did not clarify what inclusive growth is. Inclusive growth includes a cocktail of policies, which could lead to inclusive growth, but we do not know where we are heading. We cannot precisely measure inclusive growth without a precise definition, and policies do not define inclusive growth if we do not know where we are heading. We can only evaluate policies if they achieve inclusive growth, provided we know our achievement function. The following sections define inclusive growth and development, two distinct concepts.

8. Defining and measuring inclusive growth

This section provides a precise definition of inclusive growth for the first time.

The pro-poor growth is deliberately biased in favor of the poor, and its primary purpose is rapidly reducing poverty. In the previous sections, we developed a framework for pro-poor growth employing poverty social welfare functions, and these functions assign entire weight to the poor. The non-poor receive zero weight, meaning society is only concerned with the benefits of growth going to the poor and not with how the growth impacts the non-poor. In contrast, we view inclusive growth as broad-based growth, benefiting everyone, not just the poor. If the growth results in high inequality, some people receive excessive benefits, and others receive meager benefits. Recently, many countries have achieved rapid economic growth accompanied by a sharp increase in inequality, and we cannot classify such a growth process as inclusive.

Discrimination based on gender, religion, caste, or ethnicity may exclude many social groups from participating in growth. Inclusive growth ensures that all social groups can participate in economic activities and receive benefits to lead a decent life. In India, the caste system is crucial in excluding social groups such as scheduled castes and scheduled tribes from participating in the growth process. It would be challenging to link the discrimination suffered by the social groups to the inclusive growth developed in the paper. That would be our future project. The operationalizing of inclusive growth is produced below.

There is a one-to-one linkage between equality and social welfare function. How we measure equality depends on the social welfare function we choose, and we measure equality in income space using a class of social welfare functions. We view inclusive growth as broad-based growth, benefiting everyone, not just the poor. Hence, social welfare must assign positive weights to everyone's income so everyone participates in the growth process and benefits from it. The inclusive growth must also ensure that the growth benefits are equitably distributed. We cannot achieve perfect equity when everyone receives the same proportional or absolute benefits. However, inclusive growth must ensure everyone can lead a socially acceptable minimum standard of living. That means that inclusive growth must achieve higher economic growth with equity so everyone can enjoy

the minimum standard of living. It is, therefore, essential to measure the degree of inclusive growth to devise policies to enhance it.

We propose to utilize a class of inclusive social welfare functions to measure inclusive growth given by

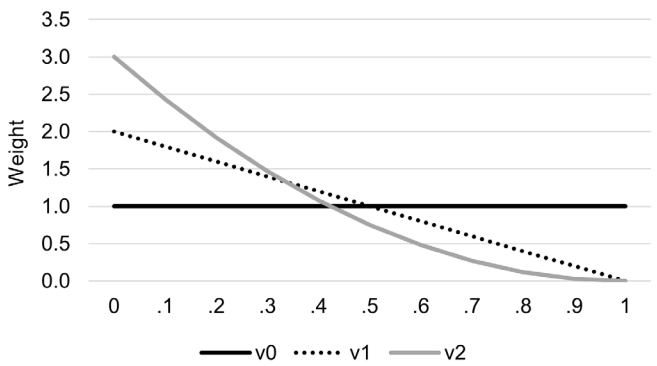
$$w(k) = (k + 1) \int_0^\infty x[1 - F(x)]^k f(x)dx. \quad (15)$$

$F(x)$ is the probability distribution function, interpreted as the proportion of persons with income less than or equal to x. The total weight given to everyone's income adds to one:

$$(k + 1) \int_0^\infty [1 - F(x)]^k f(x)dx = 1, \quad (16)$$

Kakwani and Son [2022] have proposed this general class of social welfare functions. We use this class of social welfare functions to measure inclusive growth that considers society's different value judgments. Figure 2 depicts the weighting scheme underlying the class of social welfare functions in (15). When $k = 0$, everyone in society gets a weight equal to one, in which case the social welfare $w(k)$ reduces to the average income of the society. In this scenario, society would have no concern for inequality. When $k > 0$, the social welfare function in (15) ensures that the poorest person gets the highest weight, decreasing monotonically as income increases. Hence, the richest person receives the slightest importance. This property is desirable for any social welfare function to capture income equity.

FIGURE 2. Social welfare functions weights



If $k = 1$, the social welfare function $w(k)$ reduces to the social welfare function proposed by Sen [1974]. As k increases from 1 to 2, the weight function becomes steeper, implying that the higher the value of k, the greater importance is given to the poorer person in society. k is interpreted as the inequality aversion parameter; as it increases, society gives greater importance to the incomes of the more impoverished.

Like pro-poor growth, inclusive growth can be relative and absolute. The index of relative inclusive growth is determined by

$$\delta(k) = \frac{\Delta Ln(\omega(k))}{\Delta Ln(\mu)} = \frac{\emptyset(k)}{\gamma}, \tag{17}$$

where $\emptyset(k)$ is the relative growth rate of the social welfare $w(k)$, and γ is the relative growth rate of the mean income.

If $\gamma > 0$ *and* $\delta(k) > 1$, the growth process captures the relative equity, so we define growth as relatively inclusive. The growth will not be inclusive if $\delta(k) < 1$ because the growth will not be equitable.

If $\gamma < 0$ *and* $\delta(k) < 1$, the growth will be equitable and, therefore, inclusive. If $\delta(k) > 1$, the growth will not be equitable, and hence not inclusive,

The pattern of relative inclusive growth is determined by

$$\emptyset(k) = \gamma + (\delta(k) - 1)\gamma, \tag{18}$$

which immediately shows that there will be a gain (loss) in the relative growth of social welfare if the growth process is relatively inclusive (non-inclusive).

Similar to the relative inclusive growth index in (17), we can also define an absolute inclusive growth index for the class of social welfare function $w(k)$ in (15) as

$$\delta^*(k) = \frac{\Delta(\omega(k))}{\Delta(\mu)} = \frac{\emptyset^*(k)}{\gamma_A}. \tag{19}$$

$\emptyset^*(k)$ is the absolute growth of social welfare, and γ_A the absolute growth rate of the mean income.

If $\gamma_A > 0$ *and* $\delta^*(k) > 1$, the growth captures absolute equity, so we define growth as absolute inclusive. The growth will not be absolute inclusive if $\delta^*(k) < 1$ because the growth will not be equitable.

If $\gamma_A < 0$ *and* $\delta^*(k) < 1$, the growth captures absolute equity, so we define growth as absolute inclusive. The growth will not be absolute inclusive if $\delta^*(k) > 1$ because the growth will not be equitable.

The pattern of absolute inclusive growth is determined by

$$\gamma_A^*(k) = \gamma_A + (\rho^*(k) - 1)\,\gamma_A, \tag{20}$$

which immediately shows that there will be a gain (loss) in the absolute growth of social welfare if the growth process is inclusive (non-inclusive).

9. Inclusive development

As discussed, economic growth is measured in income space, which provides people with the means to lead a better life. Means are necessary but insufficient to

give people the quality of life they must have. Inclusive development concerns the broad-based enhancement of the well-being of the population. The measurement of inclusive development requires generalizing the social welfare function given in (15). We refer to this generalization as inclusive social well-being function (ISWBF), defined as

$$\omega^*(k) = (k+1)\int_0^\infty \omega(x)[1-F(x)]^k f(x)dx, \tag{21}$$

where $\omega(x)$ is the well-being of a person with income x, when all the persons are arranged in ascending order of their income. In this function, the well-being of the poorest person in society is assigned the maximum weight of $(k+1)$, decreasing monotonically to 0 as income increases.

The relative inclusive development index for the ISWBF is given by

$$\tau(k) = \frac{\Delta Ln(\omega^*(k))}{\Delta Ln(\bar{\omega})} = \frac{\sigma(k)}{\sigma}, \tag{22}$$

where $\sigma(k)$ is the relative growth rate of social well-being, and σ is the relative growth rate of the well-being of the whole population. $\tau(k)$ captures the equity in the well-being of the society. The development will be relative inclusive (non-inclusive) if $\tau(k)$ is greater (less) than one. The pattern of pro-poor development is described by

$$\sigma(k) = \sigma + (\tau(k) - 1)\sigma, \tag{23}$$

which immediately shows that relative inclusive development leads to a gain in well-being growth rate, while non-inclusive development results in a loss in well-being growth rate.

The absolute inclusive index for the ISWBF is given by

$$\tau^*(k) = \frac{\Delta(\omega^*(k))}{\Delta(\bar{\omega})} = \frac{\sigma^*(k)}{\sigma^*}, \tag{24}$$

where $\sigma^*(k)$ is the absolute growth rate of social well-being and σ^* is the absolute growth rate of the well-being of the whole population. $\tau^*(k)$ captures the absolute equity in well-being. The development will be inclusive (non-inclusive) if $\tau^*(k)$ is greater (less) than one. The pattern of pro-poor development is described by

$$\sigma^*(k) = \sigma^* + (\tau^*(k) - 1)\sigma^*, \tag{25}$$

which immediately shows that absolute inclusive development leads to a gain in absolute well-being growth rate, while absolute non-inclusive development results in a loss in absolute well-being growth rate.

10. Pattern of global growth: a case study

This section examines whether the pattern of global growth has achieved shared prosperity and to what extent in the first two decades of the 21st century.

We have compiled the data for 173 countries over the period 2000-2021 from the World Bank's database, *World Development Indicators*, which includes four indicators:

1. Per capita GDP in 2017 purchasing power parity (PPP),
2. Life expectancy at birth in years,
3. Infant survival rate per 100 babies born,
4. Maternal survival rate per 100 childbearing women.

The pro-poor and inclusive growth is measured in income space, whereas the pro-poor and inclusive development is measured in the well-being space. Income or consumption at the individual level is ideal for measuring pro-poor and inclusive growth. However, to do so requires nationally representative household income or expenditure surveys for 173 countries over two decades, which is almost impossible.

Given this data limitation, we have carried out the analysis using each country as a unit of analysis based on countries' per capita GDP in 2017 PPP. The GDP, measured in PPP dollars, considers the prices in different countries to compare the per capita GDP across countries. The per capita GDP in the PPP exchange rate is used as a proxy for countries' economic standard of living. That enables ranking the countries from the poorest to the richest.

The main limitation of the country-level analysis is that it ignores the variations of pro-poorness and inclusiveness of individual-level growth within countries. We can only capture the between-country variations and obtain a broad picture of pro-poorness and inclusiveness globally.

In our empirical analysis, we have utilized three essential health well-being indicators: life expectancy at birth, infant survival rate, and maternal survival rate. Our analysis here is partial, ignoring many other dimensions of well-being, such as education, nutrition, living conditions, etc.

Growth rates can have wide yearly fluctuations, so it is essential to draw inferences based on trend growth rates. The least squares method applied to a semilog regression model commonly calculates the trend growth rates [World Bank 1978-2023]. Kakwani [1997] has demonstrated that it has adverse welfare implications, which are intuitively not appealing. In this section, we have used Kakwani's method to calculate trend growth rates, which satisfy all the essential properties of a social welfare function.

Table 1 presents the annual relative and absolute trend growth rates of the four indicators used in this illustration. These growth rates show that real per capita GDP has grown globally at a yearly relative growth rate of 2.20 percent between

2001 and 2020. The table also offers absolute growth rates, indicating that the real per capita global GDP has grown yearly at 286 PPP dollars. That means that the 173 countries, on average, are gaining a yearly income of 286 PPP dollars. The focus on absolute growth has an intuitive appeal because it measures the improvement in the standard of living in the income currency, easily understood by the people.

TABLE 1. Trend annual growth rates in 173 countries, 2001-2020

Indicators	Relative	Absolute
pc_gdp 2017 PPP	2.20	286
Life expectancy at birth	0.41	0.29
Infant survival rate	0.13	0.12
Maternal survival rate	0.22	0.01

That means that global growth has been significant, enhancing the average prosperity in the world in the new millennium of two decades. However, our main concern is whether this prosperity has been shared widely across all the countries, poor and non-poor (developing or developed). We answer this question by analyzing whether global economic growth has been pro-poor and inclusive.

10.1. Pro-poor global growth

Measuring pro-poor growth based on poverty social welfare functions requires identifying the poor countries. We identify a country as poor if it belongs to the bottom 40 percent of the poorest countries as determined by the per capita GDP in the 2017 PPP exchange rate. The choice of 40 percent is arbitrary; we have chosen it because the World Bank used this figure in its recently proposed development model described in Rosenblatt and McGavock [2013].

Ideally, we must construct a poverty line to measure poverty in each country based on household income and expenditure surveys. We did not follow this path because of the limited availability of the surveys.

We may turn to answer the question of whether global growth is pro-poor. We have utilized two poverty social welfare functions, one with an inequality aversion parameter of one (pswf1) and the other with two (pswf2). The higher value of inequality aversion implies greater weight to the relatively poorer countries, giving greater importance to relatively more impoverished countries.

Figure 3 answers whether the global growth has been relatively pro-poor or anti-poor. As derived, the growth is relatively pro-poor (anti-poor) if there is a gain (loss) in relative growth rate. The pswf1 contributes to a gain in the relative growth rate of 1.78 percent, whereas the pswf2 contributes to a gain in the relative growth rate of 1.23 percent. Both social welfare functions result in gains in growth rates, thus answering the question that the pattern of global growth is

relatively pro-poor. That means global growth has given more significant proportional benefits to the poorest 40 percent of the countries compared to the wealthiest sixty percent of countries.

FIGURE 3. Trend relative pro-poor growth rates

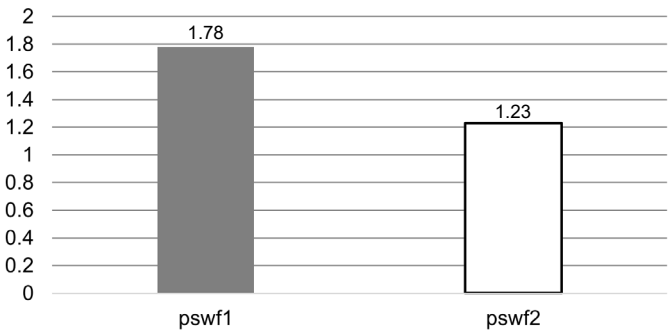
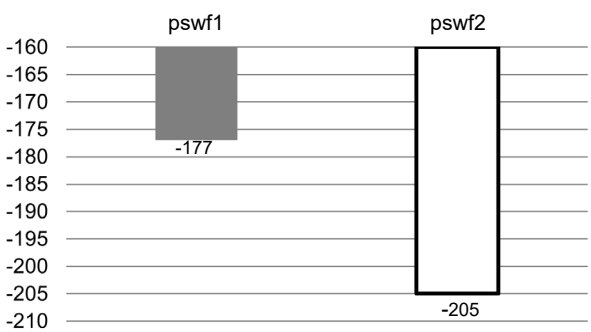


Figure 4 answers whether the global growth has been absolute pro-poor or anti-poor. As derived, the growth is absolute pro-poor (anti-poor) if a gain (loss) in absolute growth rate occurs.

FIGURE 4. Trend absolute pro-poor growth rates



The pswf1 contributes to a loss in the yearly absolute trend growth rate of USD 177, whereas the pswf2 contributes a loss in the annual growth rate of USD 205. These losses signify that the global growth rate had been anti-poor, meaning that the poorest 40 percent of the world's countries have enjoyed lower benefits of growth than the wealthiest 60 percent of countries.

The empirical conclusions emerging from this section indicate how we measure pro-poor growth matters. We have found that global growth is pro-poor in relative terms while anti-poor in absolute terms. This inconsistency may surprise many and even confuse the policymakers. However, we can intuitively explain it. Let's consider a hypothetical example of two countries, A with a per capita GDP of USD

1,000 and B with USD 5,000. Suppose A grows at a rate of 20 percent, whereas B grows at ten percent. The average growth of the two countries is 15 percent, which is relative pro-poor. The absolute growth rate of A is USD 400, and that of B is USD 500, which implies that the absolute growth is anti-poor because the more affluent country's absolute growth benefit is higher than that of the poorer country. Thus, the relative growth may be pro-poor but absolute anti-poor.

10.2. Inclusive growth

Measuring pro-poor growth requires a poverty social welfare function, which gives the entire weight to the poor; the non-poor receives zero weight. That means society is only concerned about the poor and not how the non-poor benefit from the growth. Inclusive growth is broad-based, benefiting the entire society. Therefore, its social welfare function gives a positive weight to everyone, declining monotonically as a person's income increases. Thus, even if the growth benefits everyone, to identify growth as inclusive, it must be equitable across the entire population. The growth will be non-inclusive if it is inequitable.

Figures 5 and 6 answer whether growth was relative and absolute inclusive, respectively. We answer this question based on the two inclusive social welfare functions, iswf1 and iswf2, with inequality aversion parameters 1 and 2, respectively. Figure 5 shows the gain of relative growth rates of 3.20 and 3.79 for iswf1 and iswf2, respectively. Both social welfare functions conclude that global growth is relatively inclusive.

However, the story changes when the inclusiveness of growth is measured in absolute terms. Figure 6 shows that the inclusive social welfare functions iswf1 and iswf2 result in a loss of the GDP growth rate of USD 63.24 and USD 106.72 per annum, respectively. The loss of growth rates is higher for the social welfare function with higher inequality aversion parameters, impacting the poorer countries more adversely.

Figures 5 and 6 indicate that global growth is relatively inclusive but not inclusive in absolute terms. That means how we measure inclusiveness matters.

FIGURE 5. Trend relative inclusive growth rates

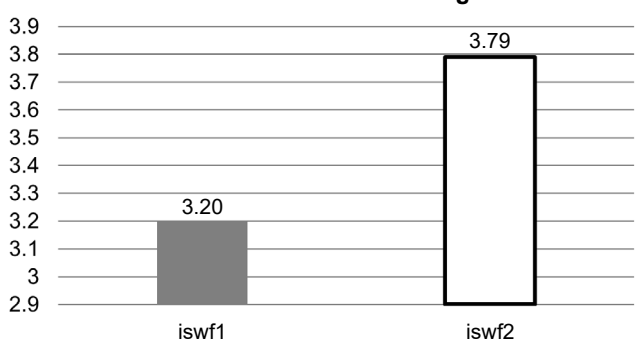
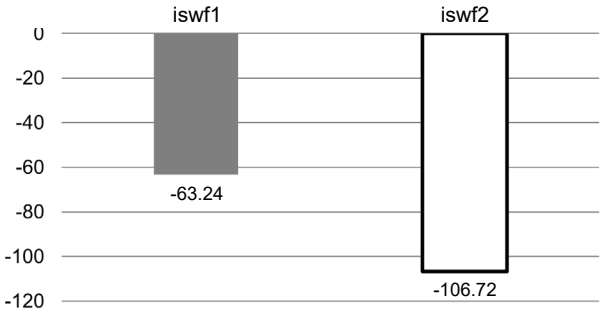


FIGURE 6. Trend absolute inclusive growth rates



10.3. Pro-poor development

How has the world performed in achieving shared development? To answer this question, we look at the pattern of development measured by the social well-being functions discussed in the paper.

Table 1 shows that the three development indicators have had positive trend growth rates, measured in relative and absolute terms. From this, we may conclude that global well-being has increased over the two decades, as measured by the three health indicators. This enhanced global well-being may have been due to the advancements in medical services that have produced improved health outcomes.

A pertinent question is whether the development has been pro-poor, benefiting the poor proportionally and absolutely more than the non-poor. Figures 7 and 8 depict the pro-poor relative and absolute development, respectively. They show that all three development indicators contribute to a gain in trend growth rates, signifying that global development was pro-poor in relative and absolute terms. That means the poorest 40 percent of countries have enhanced their health outcomes better than the wealthiest 60 percent, relatively and absolutely. That is an exciting result because it shows that the widely held perception that poor developing countries are somehow unable to progress better in their well-being compared to non-poor countries is not always true.

FIGURE 7. Trend relative pro-poor development

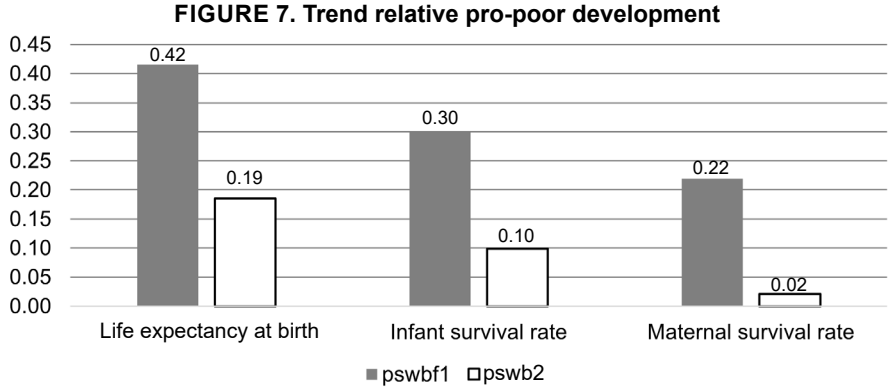
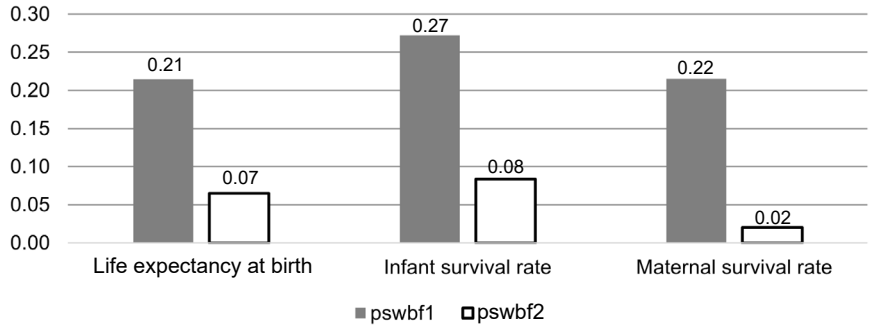


FIGURE 8. Trend absolute pro-poor development



10.4. Inclusive development

We measure inclusive development using the broad-based inclusive social well-being functions. We have utilized two inclusive social well-being functions, iswbf1, and iswbf2, with inequality aversion parameters 1 and 2, respectively.

Figures 9 and 10 show that both inclusive social well-being functions contribute relative and absolute gains in growth rates of the three well-being indicators. That leads to a fantastic conclusion that global development is inclusive in relative and absolute terms. It means the development is equitably distributed across countries.

FIGURE 9. Relative inclusive development

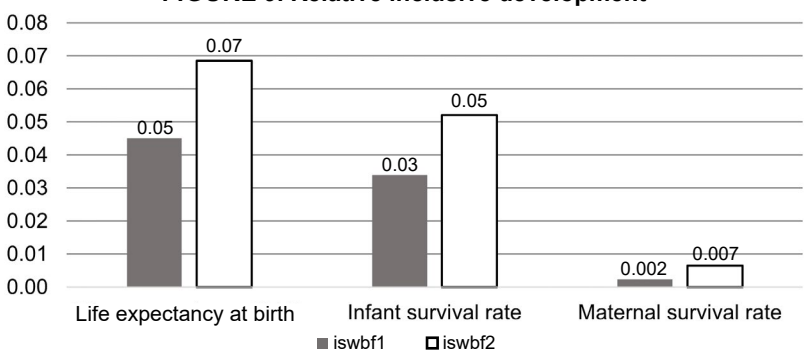
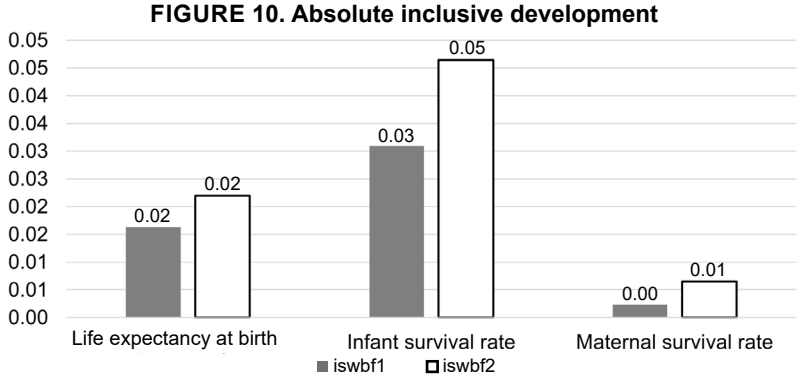


FIGURE 10. Absolute inclusive development



11. Concluding remarks

Based on a social welfare framework, the paper has developed an integrated methodology to evaluate growth and distribution simultaneously. Linking the two phenomena gives rise to four development goals: (i) pro-poor growth, (ii) inclusive growth, (iii) pro-poor development, and (iv) inclusive development. The paper has significantly contributed to defining the four concepts, providing a methodology to operationalize them using real-world data. These four development goals constitute shared prosperity, simultaneously dealing with growth and distribution. The paper has applied this methodology globally to determine whether growth and development have been pro-poor and inclusive in 173 countries over the two decades in the new millennium.

The paper concludes that global growth has been relatively pro-poor and inclusive. However, the conclusion reverses when we measure growth in absolute terms. We have explained that this inconsistency is not surprising. Even if growth is relatively pro-poor and inclusive, it may become anti-poor and non-inclusive when measured in absolute terms. That means the poorer countries have higher relative growth rates but lower absolute growth rates, which also means that the wealthier countries will tend to have lower relative growth rates but higher absolute growth rates. If this observation holds universally, we may predict that relative inequality between countries will decrease and absolute inequality will increase.

The global analysis concludes that development is both pro-poor and inclusive for the three health indicators. This finding is exciting, implying that the poorer countries have performed better in achieving better health outcomes than the more affluent countries. This conclusion may surprise many development practitioners: How can development be pro-poor and inclusive when absolute economic growth is neither pro-poor nor inclusive? We explain that this result is plausible.

As pointed out, our development concept is restricted to well-being indicators that vary in a narrow range, unlike income. For instance, the average life expectancy at birth has a maximum limit of not more than 85 years because people cannot live forever. Another essential characteristic, as articulated by Kakwani [1993b], is that achieving the same degree of improvement becomes increasingly difficult as well-being reaches progressively higher levels. For instance, it is easier to increase the average life expectancy at birth from 60 to 65 years than from 80 to 85 years. Thus, at a higher level of well-being, an incremental improvement would represent higher levels of achievement than a similar incremental improvement from a lower base. So, the relationship between achievement and values of well-being indicators is not linear.

Consequently, the observed difference in the values of indicators does not reflect the actual achievement in well-being between different countries. Thus, we must interpret pro-poor and inclusive development with caution. Kakwani [1993a] has provided a method of measuring the actual achievement of well-being indicators. Future research must utilize Kakwani's method of measuring pro-poor and inclusive development based on achieved well-being.

References

Bronfenbrenner, M. [1971] *Income distribution theory*. Routledge.

Dreze, J. and A. Sen [1989] *Hunger and public action*. Oxford: Oxford University Press.

Hanumantha Rao, C.*H*. [2010] "Inclusive growth: an overview of performance and change ahead", *The Indian Economic Journal* 58(1):3-16.

Kakwani, N. [1993a] "Performance in living standards: an international comparison", *Journal of Development Economics* 41(2):307-336.

Kakwani, N. [1993b] "Poverty and economic growth with application to Côte d'Ivoire", *Review of Income and Wealth* 39(2):121-139.

Kakwani, N. [1997] "Growth rates of per-capita income and aggregate welfare: an international comparison", *Review of Economics and Statistics* 79(2):201-211.

Kakwani, N. [2000] "On measuring poverty growth and inequality components with application to Thailand", *Journal of Quantitative Economics* 16(1):67-80.

Kakwani, N. and H.H. Son [2008] "Poverty equivalent growth rate", *Review of Income and Wealth* 54(4):643-655.

Kakwani, N. and H.H. Son [2022] *Economic inequality and poverty: facts, methods, and policies*. Oxford: Oxford University Press.

Kakwani, N. and E. Pernia [2000] "What is pro-poor growth?", *Asian Development Review* 18(1):1-16.

Klassen, S. [2010] "Measuring and monitoring inclusive growth: multiple definitions, open questions, and some constructive proposals", *ADB Sustainable Development Working Papers Series No*. 12.

Kolm, S.C. [1976] "Unequal inequalities", *Journal of Economic Theory* 12(3): 416-442.

Mahendra Dev, S. [2008] *Inclusive growth in India: agriculture, poverty, and human development*. Oxford: Oxford University Press .

OECD [2001] "Rising to the global challenge: partnership for reducing world poverty", *Statement by the DAC High-Level Meeting, 25-26 April*. Paris: OECD.

Ranieri, R. and R.A. Ramos [2013] "Inclusive growth: building up a concept", *International Policy Centre for Inclusive Growth Working Paper No*. 104.

Rauniyar, G. and R. Kanbur [2009] "Inclusive growth and inclusive development", *ADB Occasional Paper No*. 8.

Ravallion, M. and S. Chen [2003] "Measuring pro-poor growth", *Economics Letters* 78(1):93-99.

Rosenblatt, D. and T. McGavock [2013] "A note on the simple algebra of the shared prosperity indicator", *World Bank Policy Research Working Paper* 6645.

Sen, A. [1974] "Informational bases of alternative welfare approach: aggregation and income distribution", *Journal of Public Economics* 3(4):387-403.

Sen, A. [1983] "Development: which way now?", *The Economic Journal* 93(372): 745-762.

Sen, A. [2004] "Capabilities, lists, and public reason", *Feminist Economics* 10(3): 77-80

Sen, A.K. [1976] "Poverty: an ordinal approach to measurement", *Econometrica* 44(2):219-231

Sen, A.K. [2019] *Development as freedom*. New York: Alfred A. Knopf.

Sumner, A. [2003] "From Lewis to Dollar and Kraay and beyond: a review and stock of fifty years of poverty, inequality and growth", unpublished paper.

UN [2000] A better world for all. New York: United Nations.

World Bank [1978-2023] *World development report* (various years). World Bank: Washington, DC.

The Philippine Review of Economics
60(2):25-38. DOI:10.37907/2ERP3202D

Piketty inequality, meta-market failures and the new role of the state

Raul V. Fabella*
University of the Philippines
and
Asian Institute of Management

We argue that the celebrated 2014 Piketty thesis that thriving markets in advanced economies generate an ever increasing income inequality restores policy relevance to the Second Fundamental Theorem of Welfare and restores the role of the state in economics which the First Fundamental Theorem of Welfare seems, and the neoconservatives claim, to have marginalized. The Piketty thesis disproves the Kuznets hypothesis which says that the equity-deficit of market allocation is a temporary inconvenience which will dissipate as per capita income grows, thus, making state intervention unnecessary. Policies that enhance per capita growth may then replace policies of direct redistribution in the pursuit of equity. Piketty insists that this phenomenon is not due to some garden variety market failure but is due to the very dynamic that drives market prosperity, viz., private ownership of and the free enterprise deployment of capital. It is thus a meta-market failure. In properly functioning capitalist markets henceforth, the state still needs to directly push back on this meta-market failure to save capitalism from its own excesses and democracy from becoming collateral damage.

JEL classification: D31, D33, D63, 04
Keywords: Piketty inequality, role of the state, meta-market failure, fundamental theorems of welfare, well-behaved markets

1. Introduction

Income distribution and income inequality—issues beloved to Adam Smith, David Ricardo, Karl Marx, the Cambridge School, Joseph Stiglitz [1969] and John Rawls [1971]—and long been in eclipse, are back. The trigger is the book *Capital in the twenty-first century* by Thomas Piketty [2014]. Piketty, using an impressive array of empirical evidence running through two centuries and an ingenious use of barely extant tax data, tracks the trajectory of income distribution and argues that ever rising income inequality is the norm in thriving market economies. It flies in the face of the venerable Kuznets [1955] hypothesis which claimed that

* Address all correspondence to rvfabella@up.edu.ph.

in the process of development, income inequality first rises, reaches an apex and then falls as income per capita rises as it were autonomously; that the increasing inequality phase is just the imperative of a catch-up economy life cycle and not of the mature phase. In the long run in the Kuznets economy, income distribution is headed towards greater equality as economies prosper and mature. Piketty [2018] has since then written a sequel *Capital and ideology* which to his own declared taste is better, but to us seems to be only a continuation of his heroic attempt at an explanation of the runaway inequality outcome in the West; that it is ideology and fickle history, what he calls the *sacralization of property* not technology that supports, and is in turn supported by his preferred formula ($r > g$). In this new tome which extends his framework to less developed economies, he explicates his own prescription for a more equitable future for humanity, *participatory socialism*, which recommends, among others, a ponderous if well-reasoned proposal for a heavy government intervention in the form of a radical income and wealth taxation to finance a generous grant of EUR 125,000 to every citizen at age 25 to improve every poor citizen's bargaining position in the market. We will argue that this echoes the pursuit of equity in the Second Fundamental Theorem of Welfare (SFTW) and analogously in the Nash bargaining solution by redistributing initial assets which form the bargaining power (maximin position) of players. While there is no doubt that the new book ("Piketty 2018" from hereon) will only polish Piketty's claim to a Nobel Memorial Prize in the very near future, it is clearly second fiddle to his *magnum opus* of 2014 ("Piketty 2014" from hereon), which is the focus of this essay. His participatory socialism seems therefore to be more a "capitalism with socialist characteristics;" that is, a capitalism shorn of the "sacralization of capital".

At its starkest, the Piketty 2014 thesis rests on painstakingly constructed trajectories showing the income and wealth inequality (as shares of the richest income classes rather than as Gini coefficient) in income in developed countries (USA, France, Great Britain, and Europe) over the last 200 years. For example, on wealth inequality, there was a steady rise from about 1810 to a peak around 1910, followed by a marked decline until 1970 and a steady rise thereafter up to 2010.[1]To some, it seems a stretch from these trajectories to the bold Piketty conclusion that capitalism is apoptotic (see, e.g., Rogoff [2014]). It is reminiscent of the Joseph Schumpeter's response to his own query, "Can capitalism survive?" to start his influential 1942 book *Capitalism, socialism and democracy*; the response: "No." Piketty's evidence agrees with Kuznets' inequality trajectory from 1910 until 1970 and especially in the 30 years after World War II, but argues that this was an abnormal period characterized by two world wars and massive wealth destruction. Geopolitical developments after 1980 also drove inequality aversion to the sidelines: the Thatcher-Reagan Revolution, the collapse in 1989 of the Socialist Challenge to Capitalism, and Deng Xiaoping's Revolution in

[1] See, especially, Cassidy [2014].

People's Republic of China (PRC)—all happening in the 1980s and all involving a headlong embrace of the market. After all, the incentives needed for investment and growth behind a thriving market economy cannot be imagined without some inequality both as pre-condition and resultant [Chaudhuri and Ravallion 2006]. The post-1980 era became dominated by excessive inequality tolerance. The PRC in the last quarter century is exhibit A for inequality tolerance. Since poverty reduction made great strides in China in the Deng Xiaoping era in the face and most likely because of sustained rapid growth, soaring income inequality is viewed simply as a necessary inconvenience.

1.1. Inequality and growth: the development landscape

Inequality, however, stubbornly refused to quit the field figuring consistently as a factor in per capita growth regressions. Its effect on growth was predominantly negative (Alesina and Rodrik [1994]; Barro [2000]) but not in all cases (Forbes [2000]; Banerjee and Duflo [2003]), and sometimes none at all [Ravallion 2001]. Likewise, high inequality raised unemployment [Easterly 2007] and raised infant mortality (Waldmann [1992]; Tacke and Waldmann [2009]). But directly targeting greater income equity risked slowing economic growth (Ravallion [2001]; Gupta et al. [1999]) which was—and still is—viewed as the principal lever for poverty reduction (*The Economist* [2000]; Dollar and Kraay [2002]). Piketty 2018, for whom "trickle down" is a laughable excuse, of course, disagrees. On another matter it appears that the pandemic which wreaked havoc on poverty incidence in emerging economies[2] has had a modest impact on emerging markets within-country inequality measured by the Gini coefficient: a rise of 0.3 points, whereas it has been declining by about that rate per year in the last two decades. The income share of the richest one percent of income distribution declined by one percent; the share of the poorest 40 percent rose by 0.6 percent [World Bank 2022]. So the COVID-19 pandemic, which no doubt demonstrated a great resilience and forbearance of peoples to a drastically changed environment of massive threat, impacted income inequality and attitudes to it only modestly.

The first serious post-Great Recession shot for inequality aversion was fired over the sustainability of growth in the development landscape. It was widely believed that poverty reduction and the decrease in unemployment follow only when rapid growth is sustained over long periods, rather than when spasmodic [Berg et al. 2008]. But what factors contribute to sustained growth? This was the focus of the widely cited Berg and Ostry [2011] piece "Inequality and unsustainable growth: two sides of the same coin?" They asked the question: what factors prolong the duration of episodes of growth? As expected, the usual attractors of investment, namely, good institutions, openness, and stable macro

[2] The World Bank in January 2022 estimated a 0.9 percent rise in poverty incidence in 34 countries, but especially hard on poor urban households.

figured prominently as extenders of growth duration. But they also showed that income inequality shortened the duration of growth even after taking on board other factors. Concluded Berg and Ostry [2011]: "Inequality still matters, moreover, even when other determinants of growth duration—external shocks, initial income, institutional quality, openness to trade, and macroeconomic stability—are taken into account." In other words, growth in less income-equal economies tends to be spasmodic. If inequality stops growth, then the downward phase of the Kuznets curve may not even materialize. The excuse for government not to act becomes flimsy. Most of the subsequent studies involving inequality, however, still involved the less developed world where markets may exhibit myriads of distortions.

1.2. Inequality in developed economies

In the first decade of the 21st century, mounting evidence indicated that income distribution in well-functioning markets of the developed countries was becoming dangerously skewed in favor of the very rich, Piketty and Saez [2003] and Atkinson et al. [2011] revealed the sharp rise of inequality in the USA in the last two decades to levels reached in the 1920s led by stagnating real wages. Some of these may be due to financial excesses during the period. Financial institutions and regulators responding to pressures may have encouraged the process whereby high net worth individuals save while low net worth individuals borrow to indulge in unsustainable levels of spending [Rajan 2010]. The question naturally arose: was the Great Recession itself the result of extremes of financial excess and income inequality? Kumhof et al. [2015] for example linked income distribution and financial excess to the two worst economic crises in history—the Great Depression and the Great Recession.

Outside economics, Wilkinson and Pickett [2009] created a spirited methodological debate with their sweeping empirical claim that income inequality contributed to the slew of rich country social ills (infant mortality, obesity, teenage pregnancy, crime, and drug abuse, among others) in developed countries.

But as long as the good times rolled with the widely acclaimed Great Moderation, this was a matter of curiosity, not of policy. There was no dearth of plausible arguments for inequality tolerance. Welch [1999] claims that the increasing wage inequality in the US had a second edge; it also allowed greater penetration by women and men of color into the top quartiles of white men wage distribution, thus decreasing intergroup wage inequality. His rendition of the dominant attitude on rising inequality deserves repeating: "It is not much of an exaggeration to say that all of economics results from inequality. Without inequality of priorities and capabilities, there would be no trade, no specialization, and no surpluses produced by cooperation. Incidentally, there would be no economics…"

1.3. The Great Recession

Then happened the Great Recession which spread destitution among the lower income classes in the USA and other OECD countries. Median incomes of the middle and lower classes declined in the US and triggered a *zeitgeist* [Rodrik 2014] among the lower classes. The post-2008 soil had become ripe for a cloudburst on income inequality. Piketty's *Capital* hit the stands in 2014 when the USA was in the midst of a perplexing legislative gridlock over taxing the ultra-rich. His claim of ever-increasing income inequality welling out from the very engine room of progress seems to signal capitalism's inherent self-destructive tendency reminiscent of Marx's "progressive impoverishment of the masses". Its overarching message seems Marxist: the market cannot heal itself. The Piketty thesis will no doubt be subjected to further scrutiny but the criticisms thus far on the empirics have fallen flat (e.g., Giles [2014]; Krugman [2014]; R.A. [2014]). The wave of inequality awareness and aversion is welling up as a consequence. In the subsequent 2014 conference in London,[3] the IMF's Christine Lagarde and Bank of England's Mark Carney admonished businessmen to return ethics to capitalism and help push back rather than widen inequality. They did not argue how in a convincing fashion. The 2014 World Economic Forum[4] made more inclusive and sustainable growth its theme. Inequality awareness and aversion had now become front and center of policy discourse.

Section 2 of this paper explores the role of the state in properly functioning markets, the arena in which Piketty 2014 largely locates his critique and partly redressed in Piketty 2018. It argues that in properly functioning economies—those that constitute the advanced frontier economies that have led the way to whittling away at market imperfections—the state's principal role to play is to stand aside, and more so when the Kuznets hypothesis is part of the landscape.

2. The role of the state in the neo-classical economy

2.1. Fundamental theorems of welfare and the state

This section reviews the role of the state in neoclassical economics as reflected in the Fundamental Theorems of Welfare, which constitute the sanctum sanctorum of the competitive market paradigm. The First Fundamental Theorem of Welfare (FFTW), which assumes from the outset the non-existence of market imperfections and a host of convexity assumptions, shows that the market will—by itself and unaided—attain an allocation that cannot be improved upon by reallocation; in other words, Pareto efficient. The state can only do harm where no market failures exist.

[3] Conference on Inclusive Capitalism, May 27, 2014, at the Mansion House in London.
[4] World Economic Forum Annual Meeting, January 22-25, 2014, at Davos-Klosters in Switzerland.

The state is, after all, largely viewed, going back to Adam Smith, as an organ to solve market failures. Of course, it is glibly silent on the central question of who protects the property of private owners. David Hume [1739-40], for one, considered property rights as a collective action problem that requires explicit address as much as digging irrigation ditches among farmers. But if we eschew market imperfections from the start, the state has no role.

2.2. The second fundamental theorem of welfare and equity

Concern for the equity deficit in the *laissez faire* market solution motivated the Second Fundamental Theorem of Welfare (SFTW): for every feasible Pareto efficient allocation, there exists an appropriate redistribution of initial assets subsequent to which the perfectly competitive market will attain as equilibrium the very same feasible allocation. Note carefully that the statement is ingeniously an existence claim; it does not commit to a particular entity to discover and/or implement it. So the state or the government is still gingerly not in play. It could be God for all we know. But it is not lost on anyone that the state looms large as the implementor-in-waiting in the horizon. In actual asset reallocations such as in land reform programs, who but the state is the organ of execution?

As is well known, equity was championed by the then vigorous socialist challenge to capitalism (see, e.g., Lange and Taylor [1938]; Lerner [1938]) in the 1930s before Arrow and Debreu [1954] brought forth the welfare theorems. The socialist prescription to respond to the socialist calculation debate initiated by the Austrian School's anti-socialism broadside [Von Mises 1920] was "market socialism" which in time and in practice became conflated with "central planning". The *laissez faire* market equilibrium allocation may indeed and in fact be very unequal and may violate the relevant polity's sense of fairness. The main message of the SFTW is that equity and efficiency can be pursued separately. Suppose the market equilibrium allocation E violates the relevant polity's sense of fairness while another Pareto efficient allocation E^* satisfies it: should the rejection of E in favor of E^* extend to the rejection of the market mechanism? The SFTW says, "No."

If the state effects the requisite redistribution of the underlying assets, the market will attain as equilibrium the desired, more equitable allocation, E^*. The role of the state in the pursuit of equity need not go beyond initial asset redistribution; in particular, it does not have to replicate the myriad of exchanges that used to be mediated by the market using, say, the estimated shadow prices along the lines of the once celebrated but ill-fated Lerner-Lange-Taylor Theorem [Lange and Taylor 1938]. There is no need in other words for a socialized ownership of capital. This is reminiscent of Ronald Coase [1960] who argued that the assignment of property rights will, along with low transactions cost, allow private bargaining to attain Pareto efficiency in case of an externality. As in SFTW, the state enables the market to do its job by providing proper institutional mechanisms.

2.3. Meta-market failures

The SFTW suggests that another Pareto efficient allocation may be preferred by the polity to the market allocation. The competitive market model respects the ethic of Pareto efficiency as well as individual rationality. In theory, an allocation where one agent has 99.9 percent of resources while 999,000 agents have 0.1 percent of resources can be Pareto efficient. But this may violate the fairness norm of the underlying social contract upon which social stability rests. Actual societies may have norms such as fairness that transcend Pareto. This difference in norms is best reflected in different voting rights within the same democratic jurisdiction in developed economies: one-man-one-vote is the norm for political decision-making while one-share-one-vote holds in corporate decision making. The latter is an efficiency norm while the former is an inclusion norm. Economic efficiency has no mandate to be inclusive.

Take two very different allocation mechanisms: the Nash bargaining solution and the Rawlsian mechanism. The former is a two-person analogue to the market allocation; it satisfies feasibility, Pareto efficiency, and individual rationality, which the market also satisfies. The Nash bargaining solution echoes the inequality imbedded in the maximin profile reflecting the bargaining power of the bargainers and thus can be very unequal. By contrast, Rawls [1971] in *A theory of justice* argued that allocations which satisfy the "minimum difference" norm will be preferred by a polity of risk-averse members voting under the "veil of ignorance", that is, shorn of their prejudices, capabilities, and accidents of history (thus, maximin positions). The Nash Bargaining allocation will normally exhibit an equity deficit unacceptable to a polity that subscribes more or less to the Rawlsian ethic. For the two allocations to coincide requires a special property of the utility possibility frontier, sub-symmetry, which may require a state action to realize (see Fabella [1991]). If so, the Nash bargaining solution, though Pareto efficient, still constitutes a "failure" but from an equity deficit, rather than from an inability to attain the usual Pareto efficient standard. In such cases, we have what we call a *meta-market failure*: the Pareto efficient market allocation is deemed welfare-inferior to another feasible allocation on the basis of some ethical norm held sacred by the polity. This norm could but need not always be equity. But the meta-market failure stemming from the equity deficit anchored the classical objection of Marxists and socialists to the market mechanism.

2.4. Pursuing equity: the SFTW route

Let the decentralized economy R consist of two agents U and V with utility $u = \log x$ and $v = \log y$ where x and y are shares of U and V, respectively, in total resource B. The decentralized problem is that U and V must agree on a device to allocate B between themselves exhaustively. The feasible set is

$$A = \{(x, y): x + y = B\}. \quad (1)$$

Suppose the prevailing social contract prefers the allocation that maximizes the utilitarian social welfare function,

$$W = u + v = \log x + \log y. \tag{2}$$

The allocation

$$(x^*, y^*) = (B/2, B/2) \tag{3}$$

uniquely maximizes W. Suppose however that the decentralized economy R solves the problem using mechanism M which satisfies the Pareto norm but also satisfies individual rationality (respects reservation utilities or maximin positions) in pursuit of efficiency. Let $(x^\wedge, y^\wedge)$ be the allocation prescribed by M. Then it is Pareto efficient but may be deemed inferior to (x^*, y^*) from the vantage of W. Thus, R may be a meta-market failure by W.

For illustration, let the decentralized allocation mechanism M be Nash bargaining solution which chooses

$$(x^\wedge, y^\wedge) = \text{argmax}_x\, [u(x) - u^o][v(B - x) - v^o]. \tag{4}$$

Let the maximin profile be (u^o, v^o) with $u^o < v^o$. $(x^\wedge, y^\wedge)$ satisfies the first order condition

$$(B - x)x^{-1} = (\log x - u^o)(\log (B - x) - v^o)^{-1}. \tag{5}$$

Now $(B/2)$ solves this equation only if $u^o = v^o$. But $u^o < v^o$, thus, $(B/2, B/2)$ does not solve this equation and $x^\wedge < y^\wedge$. Clearly, $W(x^\wedge, y^\wedge) < W(x^*, y^*)$. Thus R is a meta-market failure. Let $D = W(x^*, y^*) - W(x^\wedge, y^\wedge) > 0$ be the potential gain from moving to (x^*, y^*).

The SFTW says that one can retain M and still attain W^* in two steps: (1) redistribute original maximin positions such that each agent gets $[(u^o + v^o)/2]$ which in the Nash bargaining game is tantamount to moving the maximin point to the 45-degree line; and (2) let M do its job. The first one echoes the Piketty 2018 proposal to raise the taxes on wealth and income to finance the 125,000-euro grant to every citizen at age 25 in order to improve the bargaining power of all citizens in the market. Since the resulting utility possibility frontier is symmetric, the resulting Nash bargained allocation is now $(B/2, B/2)$. The meta-market failure is solved by M post-initial asset redistribution. This is an analog here of the SFTW.

2.5. Transactions cost and private solutions to meta-market failures

We know that V is worse off while U is better off in (x^*, y^*) than at $(x^\wedge, y^\wedge)$. V will object to the contract involving initial asset redistribution. Consider the following side-payment contract: V agrees to the redistribution of assets in favor of U and U agrees to give up ε, $0 < \varepsilon < D$, of his/her u^* (assuming that utility is transferrable) to V such that $v^* + \varepsilon \geq v^\wedge$, and $u^* - \varepsilon \geq u^\wedge$. Thus, both are strictly better off with

the contract than at (x^, y^). If the contract is credible, and the transactions cost *TC* of decentralized bargaining is low enough, say, $D - TC > 0$, *V* will accept the contract and both parties will be better off. This is the Coasean Bargain in action. In the low-enough *TC* economy and a credible enough enforcement of contracts, the Coasean mechanism will solve the meta-market failure by private contracting. Thus, in the classical environment of small or zero *TC*, private bargaining will solve market failures but even the meta-market failure. This can explain the 'no market imperfections' assumption in the neoclassical paradigm provided there is no free riding or opportunism.

But ex-post opportunism can make for a very high *TC*, say, $TC > D$ thus eating up all the Coasean gains and the meta-market failure will remain. To exit the gridlock, one now needs a third party, say the state, with a lower *TC* to broker the deal. State intervention is itself subject to *TC*, sometimes massive. State ownership and operation of corporations can run afoul of numerous agency problems that thwart the promised maximum of consumer's surplus (see, e.g., Cook and Fabella [2002]). So many "ifs" stand in the way. But if nonetheless the state is upright and credible, it can guarantee the compensation contract among the players in the asset redistribution program. If the state intervention incurs a $TC = D/2$, the net gain of employing the state intervention is still positive. The *TC* of direct asset redistribution should never be underestimated: the program for asset redistribution may shake society to its core; social unrest can accompany such programs. The hurdle here is higher than the simple assignment of property rights in the Coase theorem, since the assets in the SFTW are already owned; whereas the common resource in the Coase theorem is, to start with, owned by nobody. Are there other ways to avoid the SFTW prescription in the pursuit of equity? The Kuznets hypothesis offers precisely such avoidance.

3. The role of the state

How big a worry the equity-based meta-market failure is depends upon what type of economy exists. In the Kuznets economy, a meta-market failure stemming from an equity deficit in the properly functioning market allocation is a benign and a temporary worry and can be accorded benign neglect. As long as over time, there is growth in income it will eventually reduce inequality. Thus, in the Kuznets economy, efforts to grow the economy which are definitely more politically feasible can conveniently substitute for SFTW and asset redistribution in the service of equity. The Kuznets hypothesis renders the SFTW policy-wise irrelevant; SFTW still remains a deep existence theorem but it is no longer a compelling policy imperative. From the Kuznetsian vantage point, the rapid rise in inequality in PRC need not worry the Chinese authorities; it will eventually abate when most of the low income households have graduated from absolute poverty into the middle class or when China finally becomes labor-scarce.

Not so in the Piketty economy: the growth in inequality respects no development threshold. Piketty [2014] rightly blames the Kuznets hypothesis for the policy neglect of income inequality; but he should also blame the Kuznets hypothesis for making the SFTW irrelevant policy-wise. If, as Piketty suggests, however, extremes of inequality threaten Armageddon on capitalism and its political mooring, democracy, then state pushback is not an option; it is an obligation. The role of the state is not only to make markets attain perfection but also to save them from runaway inequality when they do.

Piketty 2014 also echoes of Keynes' [1936] contention that the market in certain extreme conditions (due, say, to the liquidity trap or some other negative feedback process) is unable, contrary to the classical belief and to business cycle thinking started by Schumpeter [1939], to heal itself from severe systemic collapses; in such cases fiscal activism—even unorthodox ones—must come to its rescue. Fiscal activism violating prevailing common sense is what Piketty 2018 cogently argued with participatory socialism. Keynes [1924] already wrote that doing nothing is facile and unworthy: "In the long run we are all dead. Economists set themselves too easy, too useless a task if in tempestuous seasons they can only tell us that when the storm is past the ocean is flat again."

After the inequality Armageddon, the ocean may never be flat again. Or the moorings of capitalism and democracy could all be so ruptured that progress is stymied. While resemblance to Marx is apparent and even deliberate on the part of Piketty 2014, his resemblance to Keynes is also, and probably more, compelling.

4. Summary

The neoclassical market economy represented by the two fundamental theorems of welfare has hardly a place for the state. In the FFTW, the state is irrelevant because there are no market imperfections to start with. The state's role is to make perfect imperfect markets. The SFTW does recognize that the allocation attained by a properly functioning market though Pareto efficient may not meet the ethical standards of the polity, especially that of equity. In other words, the norms of society such as the norm of fairness may not be served by the competitive market which is sensitive to initial asset allocation and thus equity-blind.

When the competitive market allocation exhibits an equity deficit, we call the outcome a meta-market failure. In case of an equity deficit meta-market failure, and a more socially preferred Pareto efficient allocation that is desired over the one attained by the market, the SFTW says that there exists an initial asset redistribution that, if effected, will support the attainment of the preferred Pareto efficient allocation through the market mechanism. If the state is the organ of redistribution (over which the SFTW is discretely silent), the state does no more than redistribute the initial assets leaving subsequent reallocation to the market.

We gave an example where the social welfare function is Rawlsian while allocation is by the Nash bargaining mechanism, an analog of the market that respects both Pareto efficiency and individual rationality. The market allocation (a la Nash bargaining) will be deemed inferior by society from the vantage point of the Rawlsian social welfare function and is thus a meta-market failure. The state can intervene to correct the meta-market failure by reallocating initial assets (the maximin positions): it can craft and enforce a compensation contract whereby in the first instance the richer party gives up some of his maximin position to the poorer party in return for the poorer party compensating the richer party after the mechanism has done its job of attaining the superior social welfare level. This echoes the Piketty 2018 proposal to raise the tax on income and wealth to finance the equal inheritance scheme of EUR 125,000 euros granted every citizen at age 25. We showed that both parties benefit in the attainment of greater equity provided the transactions cost is reasonably low. The transactions cost of this program may, however, be very high (for example, it may trigger social unrest) and the state should be unwilling to push it (see e.g., Cook and Fabella [2002] for the frustration of Nirvana in the state ownership of enterprises). When, on the other hand, the state is willing to push asset redistribution, the program may fall short of realizing the Coasean dividends due to its state frailties in the form of waste and venality. The 1988 Comprehensive Land Reform Program in the Philippines which sought to redistribute land assets to landless farmers was attended with so much waste, property rights chaos and capital flight from the rural areas that the Coasean dividend headed south [Fabella 2014].

The Kuznets hypothesis offers a convenient detour. It says that over time as per capita income rises, inequality will first rise, reach a peak and then fall of itself. The meta-market failure associated with the market allocation is a transitory inconvenience in the growing Kuznets economy. In the long run, it will abate. Even when a society hankers for a more equitable distribution, it can—because costly and possibly disruptive—sidestep the SFTW; it can avoid the politically costly initial asset redistribution. Instead it can rely on trickle-down equity, that is, it can embark on policies that enable rapid income growth and wait for greater equity as a by-product. At the end of the rainbow the best of both worlds will result: high income with equity. Thus, the role of the state suggested in the SFTW is avoidable in the Kuznets economy.

Not so in the Piketty economy. Market growth induces greater and greater inequality which in turn threatens Armageddon upon capitalism and democracy. In the Piketty economy, the state is at some point forced by success to confront the SFTW; addressing the inequality-rooted meta-market failure directly is not an option but an obligation. In the Piketty economy SFTW is restored to policy relevance. The state is there not only to make the market work better but to save it when it succeeds to inequality excesses!

References

Alesina, A. and D. Rodrik [1994] "Distributive politics and economic growth", *Quarterly Journal of Economics* 109(2):465–490.

Arrow, K.J. and G. Debreu [1954] "Existence of equilibrium for a competitive economy", *Econometrica* 22:265-290.

Atkinson, A., T. Piketty, and E. Saez [2011] "Top incomes in the long-run of history", *Journal of Economic Literature* 49(1):3-71.

Banerjee, A.V. and E. Duflo [2003] "Inequality and growth: what can the data say?", *Journal of Economic Growth* 8(3):267–299.

Barro, R.J. [2000] "Inequality and growth in a panel of countries", *Journal of Economic Growth* 5(1):5–32.

Berg, A. and J.D. Ostry [2011] "Inequality and unsustainable growth: two sides of the same coin?", *IMF Staff Discussion Note* SDN/11/08. https://www.imf.org/external/pubs/ft/sdn/2011/sdn1108.pdf. Accessed July 28, 2017.

Berg, A., J.D. Ostry, and J. Zettelmeyer [2008] "What makes growth sustained?" *IMF Working Paper* 08/59. International Monetary Fund, Washington, DC. Available at https://www.imf.org/external/pubs/ft/wp/2008/wp0859.pdf. Accessed November 15, 2017.

Cassidy, J. [2014] "Piketty's inequality story in six charts", *The New Yorker*. Available at http://www.newyorker.com/news/john-cassidy/pikettys-inequality-story-in-six-charts. Accessed November 2016.

Chaudhuri, S. and M. Ravallion [2006] "Partially awakened giants: uneven growth in China and India", In L.A. Winters and S. Yusuf (eds.). *Dancing with giants: China, India and the global economy*. Washington, DC: The World Bank.

Coase, R. [1960] "The problems of social cost", *Journal of Law and Economics* 3:1-44.

Cook, P. and R. Fabella [2002] "The welfare and political economy dimensions of private vs state enterprise", *The Manchester School*. Available at https://doi.org/10.1111/1467-9957.00294. Accessed November 20, 2017.

Dollar, D. and A. Kraay [2002] "Growth is good for the poor", *Journal of Economic Growth* 7(3):195–225.

Easterly, W. [2007] "Inequality does cause underdevelopment: insights from a new instrument", *Journal of Development Economics* 84(2):755–776.

Fabella, R. [1991] "Rawlsian Nash Solutions", *Theory and Decision* 30:113-126.

Fabella, R. [2014] "Comprehensive agrarian reform program (CARP): time to let go", *Philippine Review of Economics* 51(1):1-18.

Forbes, K.J. [2000] "A reassessment of the relationship between inequality and growth", *American Economic Review* 90(4):869–887.

Giles, C. [2014] "Piketty findings undercut by errors", *Financial Times*, May 23, 2014. Available at http://webcache.googleusercontent.com/search?q=cache:http://www.ft.com/cms/s/2/e1f343ca-e281-11e3-89fd-00144feabdc0.html#axzz3ApVAKWva. Accessed August 2, 2016.

"Growth is good" [May 25, 2000] *The Economist*, http://www.economist.com/node/334693. Accessed May 2017.

Gupta, S., B. Clements, R. Gillingham, C. Schiller, M. Verhoeven, R. Alonso-Terme, and A. Mourmouras [1999] "Should equity be a goal of economic policy?", *Economic Issues* 16. Washington: International Monetary Fund.

Hume, D. [1739-40] *A treatise of human nature*. Whiteheart Cheapside: John Noon.

Keynes, J.M. [1924] *A tract on monetary reform*. London: MacMillan and Company, Ltd.

Keynes, J.M. [1936] "The general theory of money, employment and prices", http://cas2.umkc.edu/economics/people/facultypages/ ennet/courses/econ645/winter2011/generaltheory.pdf. Accessed June 20, 2017.

Krugman, P. [May 24, 2014] "Is Piketty all wrong?", *The New York Times*, http://krugman.blogs.nytimes.com/2014/05/. Accessed August 14, 2014.

Kumhof, M., R. Rancière, and P. Winant [2015] "Inequality, leverage and crises", *American Economic Review* 105(3):1217-1245.

Kuznets, S. [1955] "Economic growth and income inequality", *American Economic Review* 45(1):1-28.

Lange, O. and F. Taylor [1938] On the economic theory of socialism. Minneapolis: University of Minnesota Press.

Lerner, A. [1938] "Theory and practice of socialist economics", *Review of Economic Studies* 6(1):51-61.

Piketty, T. and E. Saez [2003] "Income inequality in the United States, 1913-1998", *Quarterly Journal of Economics* 118(1):1–39.

Piketty, T. [2014] *Capital in the twenty first century*. Cambridge, Mass.: Cambridge University Press.

Piketty, T. [2018] *Capital and ideology*. Boston: Harvard University Press.

R.A. [May 24, 2018] "Inequality: a Piketty problem?", *The Economist*, http://www.economist.com/blogs/freeexchange/2014/05/inequality-0. Accessed July 15, 2016.

Rajan, R. [2010] *Fault lines: how hidden fractures still threaten the world economy*. Princeton: Princeton University Press.

Ravallion, M. [2001] "Growth, inequality and poverty: looking beyond the averages", *World Development* 29(11):1803-1815.

Rawls, J. [1971] *A theory of justice* (Revised Edition). Cambridge, Massachussetts: The Belknap Press of Harvard University Press.

Rodrik, D. [March 13, 2014] "Piketty and the zeitgeist", *Project Syndicate*, https://www.project-syndicate.org/commentary/dani-rodrik-explains-why-capital-in-thetwenty-first-century-has-become-so-successful-in-the-us. Accessed August 5, 2016.

Rogoff, K. [May 8, 2014] "Where is the inequality problem?", *Project Syndicate*, https://www.project-syndicate.org/commentary/kenneth-rogoff-says-that-thomaspiketty-is-right-about-rich-countries--but-wrong-about-the-world. Accessed August 5, 2016.

Stiglitz, J.E. [1996] "The role of government in economic development", In M. Bruno and B. Pleskovic (eds). *Annual World Bank conference on development economics*. Washington, DC: World Bank.

Welch, F. [1999] "In defense of inequality", *American Economic Review* 89:1–17.

Wilkinson, R. and K. Pickett [2009] *The spirit level: why more equal societies almost always do better*. London: Allen Lane.

The Philippine Review of Economics
60(2):39-63. DOI:10.37907/3ERP3202D

Diamond and Dybvig in developing economies and in a digital world

Margarita Debuque-Gonzales*
Philippine Institute for Development Studies

The Nobel prize-winning article of Douglas Diamond and Philip Dybvig, entitled "Bank runs, deposit insurance, and liquidity" and published by the *Journal of Political Economy* in 1983, has spawned a large literature, including on emerging markets and developing economies. In a nod to Diamond and Dybvig, this paper reviews this subset of the literature, which has received relatively less attention than the rest despite the greater risk of banking crises in these economies; it then examines whether the seminal article remains relevant against the rapid digital transformation of financial systems today. Models that adopted their basic ideas helped drive home the importance of maintaining sound macroeconomic fundamentals and keeping confidence levels high in bank-centered economies. Similarly applying their framework to assess the impact of the current evolution of financial systems also reveals valuable insights, such as low risk from financial technology, for now, but possible shadow banks in those settings, and allows for generally better analysis, including pointing out possible blind spots when adopting new concepts, such as central-bank-issued digital currencies.

JEL classification: E02, E58, G01, G21
Keywords: Nobel prize, bank run, banking, financial intermediation, financial crises, financial fragility, liquidity crises

1. Introduction

The theoretical paper entitled "Bank runs, deposit insurance, and liquidity", written by Douglas Diamond and Philip Dybvig and published by the *Journal of Political Economy* in 1983, earned them the Nobel Prize in Economics in 2022.[1] The award came as no surprise to colleagues and "students", in classrooms worldwide, as their research had become a staple in the field of financial economics, and in the study of banking, liquidity, runs, and financial crises. Today, 40 years later, it remains, as Prescott [2010:1] put it, in the introduction to the special issue of the *Economic Quarterly* focusing on the model, a "workhorse

* Address all correspondence to mdebuquegonzales@gmail.com.
[1] Two other research papers—Diamond [1984] and Bernanke [1983]—were also recognized by the Nobel prize committee during the year.

of banking research" and "one that researchers and policymakers consistently turn to when interpreting financial and market phenomena". Yet while the general influence of the Diamond and Dybvig (DD for short) paper has been well-chronicled and examined,[2] the significance of the model for studying developing economies has received much less attention.

There are a few reasons why such a microfounded model that focuses on banks as liquidity providers, with inherent bank fragility as an offshoot of maturity transformation, would be particularly useful for analyzing less developed economies. There is, at the most basic, a greater need for liquidity in these countries than in more advanced ones, as households have relatively smaller incomes and may be more reluctant to directly fund long-term investments [Eichengreen and Rose 1998]. In the aggregate, this naturally increases dependence on transformation services offered by financial intermediaries such as banks. Banks also play a bigger role in channeling savings to investments in developing economies (as opposed to equity and debt mechanisms), so that bank runs will tend to have severe macroeconomic consequences in these settings. Constrained access of emerging markets to world capital markets compared to that of more mature economies, meanwhile, further amplifies concerns about illiquidity in the banking system, which only increases bank fragility [Chang and Velasco 2001].

However, applying the DD framework to developing countries may also be quite difficult. This would require having to consider the unique conditions of emerging market economies, where banks typically function under weaker contractual conditions and severe information asymmetries and face greater financial frictions and higher intermediation costs. Operating within small open economies and typically carrying foreign-currency debt, they also face greater challenges in managing their balance sheets because of the various mismatches (maturity and currency) and may be prone to greater risk-taking under poor regulation. More specifically, observers have expressed reservations about applying the DD framework to emerging markets, for failing to consider moral hazard and similar issues in such environments, as it assumes riskless technology.

The DD framework *purposely* excludes currency and risky technology (and hence risky bank assets) as factors in the many problems in banking, and the authors have stated that "a general model will require their introduction" [Diamond and Dybvig 1983:416].[3] DD acknowledge that introducing risky assets and moral hazard would be interesting, though hard to model, admitting that while some form of government-backed deposit insurance would remain necessary, this

[2] Coverage of the paper include the special issue of the *Economic Quarterly* published by the Federal Reserve Bank of Richmond in 2010, the scientific background paper written by the Committee for the Prize in Economic Sciences in Memory of Alfred Nobel in 2022, and numerous other papers, blog posts and web articles over the years.

[3] In an interview, Dybvig [2017] emphasized that assuming riskless assets in banks was a "feature" and not a "bug" of their model, which only served to highlight how banks tend to be (intrinsically) unstable—that is, even in the complete absence of risk.

would now have to be accompanied by some form of bank regulation to discipline risky behavior.

Despite the challenges, studies have successfully built on DD's basic model in the context of emerging markets and developing economies, considering their more complex arrangements and circumstances, and gaining deeper insight in the process. Due to the nature of the problem that the original paper addressed, subsequent research naturally focused on modeling financial crises and shaping financial regulation and policy. This paper provides a review of this interesting literature on emerging markets, and how it has helped strengthen financial regulatory and policy frameworks, written not only in a nod to the Nobel prize winners, but also to gain better understanding of these economic systems.

The line of inquiry is extended in the paper to see if the DD framework remains relevant even as the financial system transforms because of advances in digital and information technology, which have ushered in new instruments and players. The paper explores how the DD framework may provide guidance in foretelling the impact of the dramatic rise in financial technology (fintech) worldwide on financial stability. It also provides a quick assessment of how the basic DD model has figured in new theoretical models on central bank digital currencies (CBDCs), a concept that with the emergence of cryptocurrencies has become of special interest to policymakers and policy observers.

The paper is organized as follows. To provide some background, Section 2 offers a summary of the prize-winning article of DD, discusses the model's policy implications and influence, and presents a snapshot of continuing research in the area. Section 3, meanwhile, examines how DD's framework has influenced research on developing economies, including a discussion of relevant policy responses in recent crises, with reference to the work of Ben Bernanke, DD's co-winner of the 2022 Nobel Prize. Section 4 investigates the continuing relevance of DD's research under a rapidly evolving financial environment due to massive digital transformation, while Section 5 concludes.

2. DD in a nutshell (and in retrospect)

The DD paper was the first to provide a detailed framework explaining why banks exist, and why the optimal arrangement for banks—short-term liquid liabilities invested in long-term illiquid assets—also makes them inherently fragile.[4] Illiquidity of assets, according to DD, provided both the rationale for "the existence of banks *and* for their vulnerability to runs" [Diamond and Dybvig 1983:403]. As close observers have noted [He and Ma 2022], this insight from the

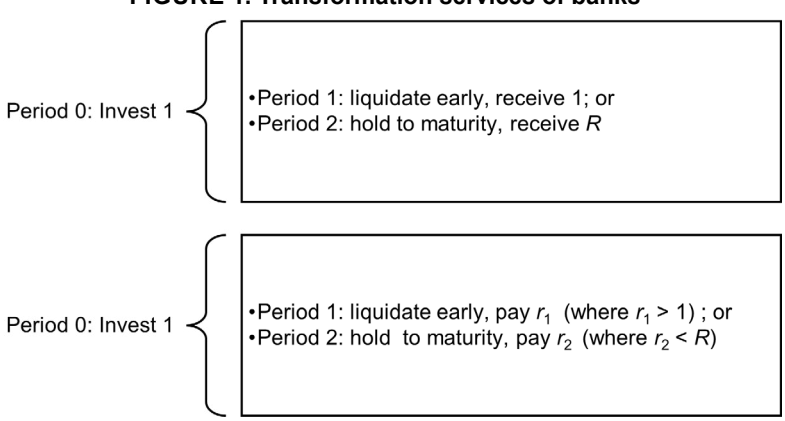
[4] Earlier research by Bryant [1980] also modeled banks as providers of liquidity and insurance against the risk of early death (generating demand for liquidity) through their offering of demand deposits. However, the paper featured a pure consumption-loans model (based on Samuelson [1958]) that did not consider productive investment and the role of banks in maturity transformation.

model would prove to be extremely valuable, as it spotlighted the critical tradeoff financial regulators face when dealing with banks, as traditionally defined, or any similar financial institution or arrangement.

In the DD model, bank assets (loans) are safe but illiquid, reflecting riskless production technology in the economy that requires time to generate returns. A simple depiction of bank assets (see Figure 1) would be that they yield R for every unit initially invested (in period 0) if held to maturity (until period 2) but just recover the investment if liquidated early (in period 1). Bank liabilities (deposits), in contrast, are short-term and more liquid, yielding more than the unit investment ($r_1 > 1$) for an investor (depositor) if instantly withdrawn but less than R ($r_2 < R$) if held for the full period. Some investors may have short-term liquidity needs in this model and will have to withdraw early, but it is not known beforehand (at period 0) who will experience such a shock.

To illustrate possible multiple (Nash) equilibria, suppose that there are 100 investors (depositors), of which only a portion will withdraw early. Without a run (the *good* equilibrium)—when only those investors who really need to withdraw their deposits on demand do so—the bank would be stable and succeed in creating liquid deposits out of illiquid loans (implying maturity transformation). However, the bank would not be able to completely service withdrawals if *all* investors decide to get their funds early (in this case, the total amount needed, r_1*100 with $r_1 > 1$, would be greater than the liquidation value of the banks' assets, which is just 100). In fact, the bank would fail even before meeting this scenario, as the number of investors who would be able to withdraw would be clearly less than 100 (again, since $r_1 > 1$). Thus, if confidence is low, a bank run may occur (the *bad* equilibrium), as investors panic and rush to withdraw their deposits ahead of the others before this critical number is reached.[5]

FIGURE 1. Transformation services of banks



Source: Author's depiction, based on Diamond [2022] and Diamond and Dybvig [1983]

[5] The demand deposit contract satisfies a sequential service constraint in the model—a first come, first served feature where depositors seeking to withdraw get paid by the bank for as long as there are remaining funds.

The DD paper shows that demand deposits, by providing liquidity, offer better risk sharing among savers with random consumption needs across periods. DD conclude that banks provide "optimal insurance contracts" where those who need to withdraw early get more than their initial deposit, while those who hold their deposits for a longer period would get a higher return, though less than what they would have received if they had invested directly in the production technology. The twist is that a bank's special function, which is to create liquidity (i.e., offer liquidity services by paying out more than the initial investment in the short term, since $r_1 > 1$) also makes it susceptible to a panic, as depositors may fail to coordinate towards the good outcome.

Briefly stated, bank fragility in the DD model essentially derives from banks' reason for existing. Moreover, since bank assets are assumed to be riskless, there is initially no solvency reason for banks to falter. This further emphasizes how damaging bank runs can be in causing "unneeded bank failures" [Diamond 2022]. Bank failures, not to mention unnecessary ones, are detrimental to the economy, as they lead to premature liquidation of long-term investments, such as through a recall of loans that interrupts production, and ultimately, output loss.

2.1. Policy implications and influence

The need to protect an economy from bank runs, specifically through government-backed deposit insurance, is the key policy implication of the DD paper (e.g., according to Dybvig in Walker [2017]). The paper argues that deposit insurance can provide superior contracts for stability in the banking system, eliminating the bad equilibrium (a run) while still allowing banks to create liquidity.[6]

Using their model, DD contrasted deposit insurance with suspension of convertibility, the other traditional device used to stop or prevent bank runs, in which deposit withdrawals are, at some point, strictly disallowed. This prevents even depositors with legitimate liquidity needs from withdrawing their deposits, making the strategy costly and unpopular.

While deposit insurance worked similarly to a central bank serving as a "lender of last resort" (LOLR) and can be modeled along the same lines, DD argued that unlike central bank lending, it is a (legally) "binding commitment" and *not* discretionary. Distinguishing between illiquidity and insolvency may be hard even for a central bank, as DD also already analytically demonstrated how even healthy banks may be pushed to insolvency because of a loss of confidence, which is quite hard to reverse in a panic.[7]

---

[6] Deposit insurance provides a guarantee, backed by the government's power to tax, that the promised return will be paid to all depositors who withdraw their funds. Note that no payout is needed if it works (i.e., if a bank run does not occur).

[7] A borrower is illiquid if it is short on cash to pay current obligations. It is insolvent if it has insufficient assets to pay outstanding debts. Central banks are normally allowed to step in as LOLR, and provide temporary liquidity, when banks are illiquid but not insolvent.

In their seminal work, DD were transparent in stating that the assumption of riskless loan portfolios (or riskless technology) precluded moral hazard problems. They confessed that such issues may exist if bank managers can choose the risk settings of loan portfolios and if this information can be kept private, or at least largely unobserved. In this scenario, deposit insurance can result in banks taking on excessive risk to bump up profits, with taxpayers left to foot the bill if bankers lose their bets.

However, DD asserted that bailouts can also introduce perverse incentives. If the LOLR always bails out banks with liquidity problems, for example, banks may make unwise gambles knowing they will be saved. If bailouts are not unconditional in the end, runs may occur with even just a shift in expectations about bank creditworthiness or about the willingness of the central bank to rescue failing institutions.

In their later work on banking theory, deposit insurance, and bank regulation, the authors argued that deposit insurance remains as "the only known effective measure to prevent runs without curtailing liquidity creation" and therefore "bank policy issues should be considered in the context of deposit insurance" [Diamond and Dybvig 1986:67]. In that paper, DD advised against the following, which they considered bad policy: limiting deposit insurance to impose market discipline on banks or requiring them to have uninsured subordinated short-term debt; using insured deposits to invest in speculative businesses that are unrelated to liquidity provision (e.g., real estate and equities underwriting); and moving towards 100-percent reserve banking, which prevents banks from fulfilling their fundamental role as liquidity provider. They also recommended tying deposit insurance premiums to the riskiness of loan portfolios, assuming risk can be somehow observed and measured, to lessen banks' incentive for excessive risk-taking.[8]

Some of the theoretical work that followed centered on extending the DD model by including risky investment choice to incorporate moral hazard effects generated by deposit insurance. These include papers by Hazlett [1997] and Cooper and Ross [2002], whose main takeaway had likewise been the importance of good regulatory design to minimize incentive problems. The latter study, for instance, finds that complete deposit insurance alone will not achieve the first best allocation and will have to be paired with additional capital regulation, as banks may opt to invest in excessively risky projects in the absence of adequate incentives for monitoring by depositors.

Over time, the fundamental policy message of DD has been interpreted to be the need for deposit insurance or access to a LOLR to avoid coordination failure and reduce vulnerability to panics, which are a natural consequence of maturity transformation [Committee for the Prize in Economic Sciences in Memory

[8] They proposed, for example, higher deposit insurance premiums for banks with many nonperforming loans, banks that previously underestimated loan losses, and banks offering above-market interest rates to gain funding.

of Alfred Nobel 2022]. Although finding the right combination of regulatory tools that will allow financial intermediaries to channel savings to productive investments while maintaining discipline in the sector remains a challenge, deposit insurance systems have already been widely established across the globe. A majority of countries have explicit deposit insurance schemes—over 80 percent of high-income countries, nearly 65 percent of middle-income countries, and about 70 percent of low-income countries—with the rest assumed to have some form of implicit deposit insurance [Anginer and Demirguc-Kunt 2018]. Meanwhile, most central banks are allowed to provide liquidity to the banking sector, albeit temporarily, to relieve frictions and avert a financial crisis.

2.2. Continuing research

Research has continued in areas considered to be the main weakness of the DD model. One pertains to the self-fulfilling nature of bank runs (or "sunspot" equilibria), which may be unrelated to economic fundamentals. In the DD model, any commonly observed random variable (e.g., a negative government forecast, a bad earnings report, or a run in another bank) can be a trigger for a panic, and this need not have anything to do with the bank's condition or the economy.

Yet empirically, banking panics have been strongly linked to weak fundamentals and the business cycle (e.g., Calomiris and Gorton [1991]; Gorton [1988]). By introducing uncertainty in the payoffs of long-term assets and the imperfectly correlated signals of these payoffs to investors, the DD model is found to produce a unique equilibrium where bank runs occur only when the expected payoffs fall below a certain threshold (e.g., Morris and Shin [2000]; Goldstein and Pauzner [2005], based on global games analysis of Carlsson and Van Damme [1993]). However, bad equilibria may still be self-fulfilling in such models, and panic-based runs may still occur even when economic fundamentals are good or when changes in these fundamentals are small. Key policy implications of the DD model therefore also remain valid.

Another key area was carved out by Jacklin [1987], who argued that introducing financial markets to the DD model would undermine banks' dominance as liquidity providers, since the social optimum could also be reached by trading in securities. This led to a study of the interplay between banks and markets. Diamond [1997] sidestepped the Jacklin critique by assuming that not everyone has access to markets and was able to show that banks and markets together can generate more liquidity than if they operate alone. This strand of research has also helped inform thinking on financial regulation, about the right amount of liquidity in the economy and particularly about the right amount of liquid reserves for banks (e.g., Farhi et al. [2009]).

The DD paper has been able to provide a solid framework for a broad spectrum of research. It has since been used to study financial contagion, sovereign debt and currency crises, and most anywhere coordination failure may play a role,

as well as various financial regulation issues. The relatively simple but logically consistent model has been able to capture the basic elements of a financial panic, which apply not just to banks, but all bank-like arrangements. These elements can be seen in almost every crisis that has occurred since the DD paper was written, and in any part of the world.

In his Nobel prize lecture, Diamond [2022] stated their findings more generally, referring to "short-term debt runs" that can bring down a solvent financial intermediary in the bad equilibrium (of multiple Nash equilibria, or multiple self-fulfilling prophesies).[9] This offered a blueprint for (private) financial crises, which he said are "everywhere and always" due to the problems of short-term debt. He pointed out that contract structure mattered a lot, as panics inherently occur if one finances long-term illiquid assets with short-term liquid liabilities.[10] Thus, runs are not limited to traditional banks; they also apply to shadow banks, a point not lost on other researchers (e.g., Prescott [2010]; Calvo [2012]; Adrian and Ashcraft [2016]), who have cited the role of runs on such institutions in the Global Financial Crisis (GFC) of 2008/2009.[11]

3. DD in developing economies

This section summarizes the important research adapting the DD framework to emerging-market settings, which mostly focus on financial crises, to gather insights from the literature.[12] It also discusses the implications of this line of research for financial regulation and policy, and how the DD model influenced policy responses in recent crises, including the COVID-19 pandemic.

The latter discussion cites the contribution of Ben Bernanke, the other Nobel laureate for economics in 2022, who provided the rigorous empirical analysis needed to show how bank failures and a credit crunch could leave deep economic scars [Bernanke 1983]. In the policy realm, as similarly observed by Kashyap [2015], he may be cited for recognizing the basic elements of a crisis, as sketched out by the DD model, apart from sharing his vast knowledge from his own research, prompting policymakers (including himself) to quickly address panics even among bank-like institutions during the GFC.[13]

---

[9] The authors acknowledged this in the original paper—i.e., that the potential for multiple equilibria did not apply solely to banks. They chose to focus on banks, as banks were known to account for a large portion of corporations' short-term debt.

[10] A panic may occur if creditors start to lose confidence in the borrower and pull out their funds or if they worry that other creditors would respond that way.

[11] Shadow banks refer to non-depository institutions that engage in maturity transformation but are not subject to traditional bank regulation. They include the different types of funds (e.g., structured investment, hedge, and money market mutual funds), securities companies, and consumer finance institutions.

[12] This is not meant to be an exhaustive literature review. Rather, the aim is to get a flavor of the research that the DD framework has spawned in the context of emerging markets and developing economies.

[13] Ben Bernanke served as the Chair of the US Federal Reserve during 2006-2014.

3.1. Modeling financial crises

The DD paper has been pivotal in the context of emerging markets by providing the workhorse model that enabled the development of liquidity and bank-run models of financial crises during the 1990s (Masson [1999]; Frankel [2010]). Related studies were mostly motivated by a series of emerging market crises that marked the period, such as the Mexican Crisis in Latin America in 1994 and the Asian Financial Crisis (AFC) in 1997/1998.[14]

One of the earlier papers that surfaced in that era was written by Sachs [1996], who argued that one possible cause of financial crisis in emerging markets was a self-fulfilling panic, the most familiar case being the self-fulfilling *banking* panic outlined by the DD model. Other known triggers then included exogenous shocks to markets, inadvertent policy shocks, and exhaustion of borrowing limits.

Subsequent research in the area included the study of Radelet and Sachs [1998], who depicted the crisis in Mexico as essentially a creditor run on government debt, while characterizing that in Argentina, which occurred a year later, as a creditor run on the banking system. The authors further suggested that the AFC, was similarly, in essence, an international variant of a bank run—this time, with international bank debts owed mainly by the financial and nonfinancial corporate sectors (in Korea and Thailand, and Indonesia, respectively).

DD helped set the theoretical basis for their analysis, with the Asian crisis analyzed as being, in large measure, due to self-fulfilling tendencies in the financial system. Radelet and Sachs [1998] noted the DD model offered a "much more complete theory" to explain self-fulfilling panics (i.e., crises with multiple rational equilibria) in the context of banking institutions than its precursors. Presumably guided by this literature, they claimed that it had been the refusal of foreign lenders to roll over short-term credit, rather than any fundamental weakness in Asian economies, that triggered the AFC.

Like Radelet and his coauthors, Chang and Velasco [1998;2001] attempted to reinterpret the financial crises of the 1990s as international versions of a bank run. The authors made theoretical advances by developing an open economy version of the DD framework, which focused on the microeconomics of banking, allowing them to formally model financial fragility in emerging market economies. Their version basically embedded banks in a small open economy. The main departure from the original framework was that it allowed access to international capital markets, with a domestic bank permitted to borrow abroad, to help fund both long-term investment and immediate withdrawals.

[14] Other explanations for the financial crises in emerging market economies during the period are of course available, though we do not discuss them in here. Other well-known papers on so-called third-generation models motivated especially by the AFC emphasized the role of: hidden subsidies and moral-hazard lending resulting in overborrowing (e.g., McKinnon and Pill [1997]; Dooley [2000]; Burnside et al. [2004]), and corporate balance sheet effects and capital flows and their impact on the real exchange rate [Krugman 1999].

With this twist in DD's banking story, Chang and Velasco concluded that a bank run may occur in an emerging market when a domestic bank experiences *international* illiquidity. While fundamental weakness underlies crises in this model, foreign creditor panics, which can occur if local banks cannot commit to preserve enough resources for foreign debt payments, can trigger runs on domestic deposits in their model (and vice versa), with the likelihood depending on the maturity of foreign debt and the possibility of international default. Financial liberalization and short-term capital inflows can worsen bank illiquidity and increase financial vulnerability in the model. Their theoretical findings generally appeared to match empirical observations in Latin America and Asia during the period, when short-term foreign debt notably increased financial fragility in some countries by heightening rollover risk.[15]

Research on financial crises in developing countries seemingly stalled in the 2000s, as the epicenter shifted to advanced economies, owing to the suddenness and severity of the GFC in 2007/2008. New studies have since emerged, lengthening the research thread that made use of the DD model to analyze developing-economy crises or as a building block for macroeconomic models designed after emerging markets.

While not concentrating on developing economies, Calvo [2012] built a model that could explain some of the central stylized facts not just of the subprime crisis in the US, which spread to other global markets during the GFC, but also of the sudden stops and previous emerging market crises, with the DD model's intuition at its core. In the model, which encompasses bank-like arrangements such as "shadow banks", financial development is cast as a mechanism that endows real assets (such as land and capital) with liquidity, which may be impaired by shocks that are analytically equivalent to a bank run. This setup allows for bubble-like episodes that are not driven by fundamentals but may be fully rational. Although Calvo [2012] deemed the model to be still highly incomplete, he argued that it nevertheless bared new insights about the effects of liquidity creation and destruction.[16]

To study banks and liquidity crises in emerging market economies, Matsuoka [2018] later extended the model of Chang and Velasco [2001] by incorporating interbank asset markets, as well as the models of Allen and Gale [1998;2004a;2004b], which are also based on a DD economy but with aggregate shocks and aggregate uncertainty introduced to the system. The resulting banking model generates two types of equilibrium: a no-default equilibrium and a mixed equilibrium. In the latter, *risky* banks default, while *safe* banks meet their commitments and ultimately purchase the long-term assets of risky banks. Matsuoka [2018] states that the model generally succeeds in capturing the basic

[15] Furman and Stiglitz [1998] remarked that the ratio of short-term debt to reserves, by itself, was able to predict the East Asian crisis.

[16] It can show, for instance, the possibility of "excessive" financial innovation, where an increase in capital liquidity may lower individual welfare. An extension of the basic framework, meanwhile, finds support for the conjecture that low policy interest rates may have provided further incentives to "shadow banks".

features of banking crises in emerging market economies (e.g., internationally illiquid domestic banks and bank assets traded at fire-sale prices), particularly after financial liberalization, where large capital inflows intensify asset price volatility and exacerbate banking crises.

3.2. Shaping financial regulation and policy

The DD framework highlighted the importance of having deposit insurance and a LOLR, combined with bank regulation that reduces moral hazard, to avoid a financial crisis. Expanding the DD model to better match conditions and capture issues in emerging market economies, as chronicled above, helped point to the need for further refinements in financial regulation and policy in these countries.

Framing financial crises in emerging markets as international versions of a bank run held important policy implications, especially for managing capital flows and regulating banks and other domestic financial institutions.[17] It underscored, for instance, the wisdom of avoiding a buildup of short-term debt—particularly short-term foreign debt—which could increase the likelihood of coordination failures among creditors, heightening vulnerability to runs (Chang and Velasco [2000;2001]).[18] It also reinforced the argument for better management of capital account liberalization and capital flows, precisely referring to large loan volumes contracted at short maturities and in foreign currency, which can worsen maturity and currency mismatches and bank illiquidity.

Considering the self-fulfilling nature of panics, a major policy lesson supported by bank-run models revolved around the avoidance of triggers that could lower the confidence of creditors and coordinate them towards a bad equilibrium. As past emerging market crises featured undercapitalized banks, greater attention was placed on strengthening regulation and supervision of banks and raising capital adequacy standards (e.g., Sachs [1996]; Radelet and Sachs [1998]; Roubini [2000]).[19]

In Asia after the AFC, for instance, major reforms included cleanup of nonperforming loans of banks (through special purpose vehicles and other bad-debt resolution mechanisms), stronger macroprudential policies, more intensive bank monitoring, establishment of better risk management practices and prudential controls among banks, and the drive for greater accountability and

[17] It should be mentioned that the possibility of international bank runs also provides a rationale for an international LOLR. This paper does not discuss such issues, though they are clearly important, but focuses instead on domestic-level policies over which local policymakers have some control.

[18] The oft cited and logical policy solution was to lengthen debt maturities. However, this was not without caveats. Some argued that short-term debt fulfills a function by serving as a commitment device for the borrower (Rodrik and Velasco [1999]; Jeanne [2009]), while others contended that shifting toward longer debt maturities may have a destabilizing effect on the banking system [Matsuoka 2018].

[19] Commonly cited proposals included proper sequencing of financial liberalization reforms, with steps to strengthen regulation of the financial system coming first, and during the time, taxation to slow down capital flows (such as that used by Chile, which imposed a 30-percent reserve requirement on dollar deposits in the banking system), respectively.

transparency in corporate boards.[20] Deposit insurance was also adopted around the 2000s in many economies in the region.[21] Altogether, these efforts helped raise public confidence in the region's banking systems and protect Asian economies from the harsh effects of succeeding global crises.

Other possible policy-related triggers of banking crises in emerging markets that were often cited included overly expansionary monetary and fiscal policies that could fuel a lending boom and fixed exchange rates [Eichengreen and Rose 1998]. The latter made a bank run more likely as it prevents the central bank from acting as a LOLR when needed, as doing so (providing liquidity to distressed banks) may threaten the currency peg.[22] As emphasized in the DD paper, credibility as well as capacity of authorities was crucial to maintaining bank stability, whether in the case of deposit insurance or a LOLR. Based on this analysis, preserving confidence and reducing vulnerability to financial panics would require having sound macroeconomic fundamentals, notably including greater exchange rate flexibility and a healthy level of foreign reserves, particularly where short-maturity foreign debt may be a concern.

While deposit insurance has been taken as the standard policy recommendation of the DD paper, such systems have not always worked exactly as intended, especially in poor institutional environments. Under weak settings (such as uncertain rule of law and widespread corruption) and if poorly designed, they may serve to erode market discipline, destabilize the banking system, and hinder growth and financial development (see Anginer and Demirguc-Kunt [2018] for a discussion of the empirical findings). Increasing attention has therefore been placed on improving the design of these systems, specifically by incorporating features that internalize risk-taking by banks to reduce moral hazard.

3.3. Policy responses to more recent crises—shoutout to Bernanke

While DD provided the theoretical explanation for the existence of banks, their vulnerability to runs, and the damaging nature of runs, it was Ben Bernanke, their fellow Nobel laureate in 2022, who provided the empirical evidence on the problem. In his prize-winning empirical research on the Great Depression, bank failures were revealed to be largely responsible for the exceptional depth and duration of that historical downturn. DD cited Bernanke [1983] in their paper, which was written during the same year, emphasizing how Bernanke's research

[20] Other explanations for the Asian crisis—such as moral hazard lending and overborrowing, including by related parties, which represented hidden deficits—also helped encourage reforms in this area. There had been a strong push to lessen the dependence of Asian financial systems on banks and the implicit guarantees offered by governments to these financial institutions, particularly through the development of local-currency bond markets.

[21] Explicit deposit insurance was unavailable in the original ASEAN-5 during the AFC, except for the Philippines [Noman et al. 2022]. It was soon introduced in Indonesia (in 2004), Malaysia (2005), Thailand (2008), and Singapore (2010).

[22] That is, a bank run may spur a run on the domestic currency if the central bank tries to fulfill this role [Masson 1999].

supported their thesis by showing that bank runs were indeed a better predictor of distress than money supply.

This would seemingly be mirrored a couple of decades later. In a speech he made as US Federal Reserve Chair in 2009, Bernanke noted that while economic fundamentals played a role in triggering the GFC or Great Recession, the ongoing crisis also exhibited features of "a classic panic", which the DD paper was able to break down. In 2018, he provided further empirical evidence that the Great Recession was primarily due to a financial panic in funding and securitization markets (essentially shadow banks engaged in maturity transformation), which eventually spread and led to a disruption of credit supply.[23] He argued in that article that this finding helped "justify the [US] government's extraordinary efforts to stem the panic in order to avoid greater damage to the real economy" [Bernanke 2018:251].

Following the lead of the US Federal Reserve and the European Central Bank—with the US Fed acting as LOLR of both traditional and shadow banks following the collapse of Lehman Brothers—other central banks intervened to avoid short-term debt runs and preserve credit supply. In each country, the end goal was to avoid a deep and lasting recession. The Bank of Japan (BOJ) also took steps to secure the stability of the country's financial system, including stock purchases from and provision of subordinated loans to banks, and to facilitate corporate financing [Bank of Japan 2023]. Remarkably, Asian banking systems, which had undergone a regulatory and policy overhaul after the AFC, were then much less exposed to US subprime assets. Although output in Asia also contracted during the period, it did so to a lesser degree, and economies recovered much faster than the rest of the world [Jeasakul et al. 2014].

With the relative success of these strategies during the GFC, similar interventions were applied during the COVID-19 pandemic, and on a bigger scale in many countries given the nature of the crisis. Public health concerns then led to the closure of contact-intensive sectors of the economy, leading to large drops in output and high unemployment. The pandemic had been loosely interpreted as a natural disaster that froze economic activity—and the ensuing economic crisis as *not* being due to bad fundamentals such as fiscal recklessness or excessive financial risk-taking. Therefore, emphasis was placed on protection, and not punishment, as had been deemed warranted to prevent moral hazard in the light of past (financial) crises.

Amid the uncertainty of the COVID-19 pandemic, central banks again sought to ensure continued flow of credit to the economy and prevent a credit squeeze from developing into a full-blown financial crisis that could fuel a deeper recession. Several measures were again taken to supply liquidity to financial institutions

[23] Runs in different markets were recorded during the GFC, including in asset-backed commercial paper [Covitz et al. 2013], money market mutual funds [Schmidt et al. 2016], and the repo market [Gorton and Metrick 2009]. See also Prescott [2010].

(e.g., by lending to these institutions, purchasing their assets, or switching their illiquid assets with more liquid securities) so they in turn could help firms and households, especially the smaller or weaker ones, weather the pandemic.[24] Authorities in many countries also displayed regulatory forbearance during the COVID-19 crisis, particularly for banks, in a bid to further ease credit conditions.

The BOJ, for example, introduced a new "funds-supplying" measure that provided liquidity to private financial institutions, in substantial amount and on favorable terms, to facilitate lending to small and medium-sized firms [Kuroda 2020]. Meanwhile, among the developing economies in Asia that experienced the AFC, the central banks of Malaysia, the Philippines, and Thailand likewise launched measures that supported lending to smaller enterprises [IMF 2021]. Similarly, the central bank of Mexico, which also suffered a crisis in the 1990s, was able to open financing facilities for banks allowing them to channel funds to micro, small, and medium-sized enterprises and individuals affected by lockdowns. Indonesian authorities provided regulatory relief to domestic banks, as did most of its neighbors in the developing ASEAN region. In addition, the Bank of Indonesia was allowed by presidential decree to finance the country's deposit insurance agency through repo transactions and purchases of government bonds owned by the agency.

Hence, policymakers apparently still take a leaf from the DD paper, including in developing economies. Didier et al. [2021] observed though that banks did not experience major liquidity problems during the pandemic crisis, unlike in a typical financial crisis. Instead, there appeared to be limited appetite, as firms faced heightened credit risk on account of the uncertain nature and path of the COVID-19 virus. Central banks worldwide had difficulty in this regard, as extending liquidity lines and similar policies worked only if the funds were indeed passed on to and utilized by firms.

The DD framework nevertheless suggests that such policies to preserve credit supply still had great merit, as the alternate scenario may have been far worse. A key feature of the model, after all, was the critical role of confidence and credibility in avoiding a bad equilibrium (a run) for banks and bank-like arrangements. Such stability, in turn, is needed for a smooth and steady functioning of the real economy.

4. DD in a digital world

Finally, it would be interesting (and useful) to ask if it is likely that the DD paper, which looked at the microeconomics of "banking", will remain relevant even as the financial system evolves and reveals new actors. While we know this

[24] Brunnermeier and Krishnamurthy [2020] argued that policy should focus on the survival of viable firms, and advocate for a pause (not bankruptcy as was the policy in past crises), particularly for small and medium-sized enterprises (SMEs). Since SMEs were less able than larger firms to withstand a liquidity crunch, they would benefit from ample provision of low-cost refinancing for rolled-over loans to stabilize existing businesses. The authors stated that the more beneficial goal was to "evergreen" the SME loans until the pandemic subsided.

is likely to be so, how exactly can the policy lessons from the DD framework help authorities navigate the fast-changing financial intermediation landscape? In which areas of the fintech environment can it provide illumination and guidance? This section focuses on the development of fintech to provide at least some (partial or first-pass) answers to these questions. It also provides a glimpse of how the DD framework has figured in the analyses of CBDCs, which emerged after the advent of cryptocurrencies to become a much-debated topic.

4.1. Rise in fintech

Recent years have seen rapid technological change in the financial industry, with the rise in fintech accelerating during the COVID-19 pandemic, which stimulated demand for digital services and the adoption of digital finance. Over time, fintech has evolved from traditional financial institutions simply using information technology (IT) to improve products and services (e.g., electronic payments and clearing systems, ATMs, and online banking) to the entry of new players, also aided by IT, providing *non-intermediated* financial services directly to customers and creating a whole new environment for financial institutions [Thakor 2020]. A wide variety of online models soon started to compete with traditional "brick-and-mortar" banks in key areas such as payments, remittances, and lending, among others [Murinde et al. 2022].

Fintech's fluid development has made it hard to classify, prompting broad definitions. The Financial Stability Board, for example, defines fintech as "technology-enabled innovation in financial services that could result in new business models, applications, processes, or products with an associated material effect on financial markets and institutions and the provision of financial services". The Basel Committee on Banking Supervision (BCBS), meanwhile, has categorized fintech innovations into three broad product sectors—namely, credit, deposit, and capital-raising services; payments, clearing, and settlement services; and investment management services—in addition to market support services [BCBS 2018].

The DD framework tells us that financial fragility exists whenever illiquid assets are financed by short-term debt (or whenever transformation services are offered). It may therefore shed light wherever borrowing and lending (or financial intermediation) occurs, such as in the credit, deposit, and capital-raising space of fintech.

There are two concerns about this rapidly changing area. The first relates to the impact of fintech players on incumbent banks, and the possible effects on financial stability, as the former may exert competitive pressure on the latter, pushing them to take greater risks to recover their profits, or replace them completely. The second pertains to the vulnerability of the fintech players themselves.

As the DD model has helped crack the code on financial fragility, one can look at the key assumptions and mechanisms and see if they are present in the problem at hand. For instance, comparing the differences between banks and peer-to-peer

lending (P2P) platforms, currently the largest form of fintech financing [Bollaert et al. 2021], can be quite informative. Based on such an exercise, Table 1 from Thakor [2020] seems to suggest that one need not worry so much (yet) about the latter's impact on financial stability.[25]

There are many reasons from the literature explaining why fintech lenders may not eliminate banks.[26] From the DD model, the most crucial would be that they do not offer many of the services provided by banks, particularly risk sharing, liquidity creation, and consumption insurance.

Without a banking license, fintech lenders can only raise the necessary funds but cannot offer transformation services [Navaretti et al. 2018]. Acting like brokers in an agency model, they match counterparties and receive fees for this service, but they cannot use the pooled funds to finance illiquid loans or less liquid assets. They are therefore much like "narrow" or full-reserve banks, which DD declared provided no liquidity services [Diamond and Dybvig 1986]. As they no longer hold credit risk on their balance sheets, fintech lenders also do not function as "delegated monitors", another important function of banks established by Diamond [1984], in his other prize-winning paper.[27] As equity holders, the investors will have to do the credit monitoring and collect the required information themselves.

The DD model tells us, however, that it is for the same set of reasons why bank-like fragilities may be less of a concern in the current fintech environment. The FSB [2017] observes that most P2P lending platforms are unleveraged, unlike banks, with only a small proportion of platforms using their balance sheets to fund loans. Moreover, the P2P lending model does not entail bank-like liquidity risks, as investments and loans are typically duration-matched. An investor may not liquidate their investment before maturity date and will need to find another investor to take their place before they can exit.

However, fintech lenders are more vulnerable than banks to operational risks (such as cyber risks, disruptions to outsourced IT services, and fraud-related risks, including money laundering and corporate misconduct) and misaligned incentives under an agency lending model adopted by most platforms. The FSB [2017] reports that the business models of these online lenders are more like the "originate-to-distribute" model of mortgage lenders prior to the GFC, indicating moral hazard risks, especially if higher fees are charged to higher-risk borrowers (such as with fees set proportional to interest rates) or to investors (upon loan collection).

[25] In P2P lending, P2P platforms, after preliminary credit analysis, combine loan bids by investors into loans but do not invest in these loans. Funds provided through these platforms are therefore closer to investor equity, with P2P lending considered as *non-intermediated* finance.

[26] These are apart from the regulatory advantage in terms of deposit insurance and LOLR guarantees that give banks a funding-cost advantage and allow them to meet investors' demand for safe assets.

[27] Simply stated, the theory implies that banks can obtain funding even for high-return but risky projects, as they can commit to pay their creditors (depositors) by monitoring borrowers on their behalf and through diversification of their loan portfolios.

TABLE 1. Banks vs. fintech lending (peer-to-peer platforms)

Banks	P2P lending platforms
A. Services provided	
• Improved risk sharing and consumption insurance	• No
• Screening	• Yes
• Monitoring	• No
• Funding liquidity creation	• No
• Loan commitments (credit rationing insurance) and other off-balance-sheet puts and guarantees	• No
B. Capital structure	
• High leverage with little bank equity capital	• All equity-financed: no equity capital invested by lending platform, so investors are equity holders in loans
C. Incentive problems	
• Insufficient screening	• Yes
• Insufficient monitoring	• No
• Insufficient funding liquidity creation	• No
• Excessive risk-taking due to high leverage and safety nets	• No
• Overlending and excessive growth due to incentives distorted by safety nets and too little capital	• Overlending and excessive growth due to profit-maximization motives.
• Insufficient capital due to safety nets	• No
• Incentives to renege on off-balance-sheet commitments	• No
D. Regulation	
• Deposit insurance and capital regulation	• No
• High regulatory costs and restrictions	• Lower regulatory burden
E. Objective function	
• Maximize bank equity value	• Maximize value of P2P platform's owners' claim consisting of origination and other fees plus fraction of borrower repayments

Source: Thakor [2020].

Many studies have correspondingly argued that fintech will not replace traditional finance any time soon. Based on a review of the literature, Thakor [2020] concludes that P2P lending may take some market share away from banks but will not replace bank lending "in the near future", with P2P lenders likely to take risky borrowers (those who lack collateral) away from capital-constrained banks. He expects banks to eventually build their own online lending platforms, acquire such platforms, or partner with existing platforms. Murinde et al. [2022],

Bollaert et al. [2021], and Navaretti et al. [2018] similarly share the view that fintech lenders are unlikely to supplant banks, but may coexist with them, cooperate with them, and/or evolve together to become more like each other (e.g., banks developing their own fintech platforms or working with fintech startups, and fintech lenders possibly engaging in maturity transformation to some extent, to provide greater liquidity services).

Yet that is as far as the benign evolution of the financial system goes. Based on lessons from the DD framework, there are two areas that may need to be watched. One connects to the existence and possible spread of shadow banks in the fintech lending space—financial intermediaries that are neither P2P lending platforms, though they use IT extensively in lending, nor banks, despite having similar balance sheets [Thakor 2020]. Like banks, they perform liquidity transformation and invest their own equity, but unlike banks, they obtain funding through uninsured debt financing or via securitization instead of deposits. As shadow banks, they are unregulated and inherently fragile entities that may be susceptible to runs, as had been the painful experience during the GFC.

The other area to watch refers to the potential for systemic importance of aggregators in finance, as they may become the default solution for accessing banks, when applying for new accounts and loans [FSB 2017]. Some now instantly link to digital banks or neobanks—a fintech innovation that shifts away from relationship banking—and not just to online versions of traditional banks.[28] While such an arrangement may improve financial inclusion as hoped, it may also create new risks, as loans and deposits become more sensitive to financial and real shocks [Gambacorta 2023]. Whether or not it will worsen financial fragility remains to be seen. From Diamond [1984], we recognize that much depends on how well loans may be selected, diversified, and monitored even without human interaction. From DD, we know that it will hinge on how confident (and trustful) depositors and other investors may turn out to be in such arrangements.

4.2. Emergence of CBDCs

As a final illustration, this subsection briefly notes how the DD model has contributed to the theoretical analysis of CBDCs, an idea spurred by the development of distributed ledger technology, which enabled decentralized settlement of electronic transactions and the creation of cryptocurrencies. CBDCs possibly eliminate the need for physical cash; allow the central bank to engage in large-scale intermediation, in competition with private banks for deposits and likely involving some form of lending of those deposits; and, in summary, permit consumers to directly hold a bank account with the central bank [Fernández-Villaverde et al. 2021].

[28] Digital banks, unlike their traditional counterparts, rely on a business model that is based mainly on technology and data. They have no "brick-and-mortar" facilities and rely solely on mobile phones and apps (i.e., no human interaction). To reduce the need for collateral, they make use of machine learning and non-traditional data as inputs to credit scoring.

CBDCs can improve welfare by reducing frictions in deposit markets and payments, encourage financial inclusion, and improve the transmission of monetary policy (Infante et al. [2022]; Ahnert et al. [2022]). However, they can also carry risks, such as possible bank disintermediation, as they may increase the funding cost of banks and reduce bank lending; and potentially greater bank fragility and higher likelihood of systemwide runs.

The DD model, considered to be the canonical model of bank runs, has proven vital in studying the potential impact of CBDCs, particularly on financial stability. The new models, in turn, have been informative especially for policymakers, who are also just grappling with the concept. Fernández-Villaverde et al. [2021], for instance, introduce a central bank and a CBDC to the seminal model, allowing them to investigate the implications of a CBDC account that potentially competes with traditional deposits in commercial banks.[29] Unlike commercial banks, a central bank can only invest in long-term projects through investment banks. In addition, it cannot terminate these projects prematurely (wholesale loans to investment banks are not callable and hence protected from early liquidation), and it can default without going bankrupt. There is therefore little incentive to run on the central bank in the resulting model, while commercial banks remain fragile for reasons outlined in the DD paper.

The authors further argue that the rigidity of the central bank's contract with investment banks will eliminate the run equilibrium, making the central bank more stable than the commercial banking sector. Realizing this, consumers choose to deposit exclusively with the central bank, and the latter becomes the "monopoly provider of deposits" in the economy, which possibly "endangers maturity transformation". This arrangement could jeopardize the independence of the central bank, which now has the power to invest in specific projects and may face political pressure as a result.

In a related paper, Schilling et al. [2020] create a nominal version of the DD model. It differs from the classic setup by considering central bank intermediation, which allows the monetary authority to control the price level. CBDC accounts are now nominal rather than real assets or claims. CBDC depositors may "run" on the central bank by rushing to spend their nominal balances, with such behavior possibly triggered by concerns that their holdings may start to lose value. In this model, real value is determined by the central bank's liquidation policy for its real investment, where selling more of the (illiquid) asset would place downward pressure on prices, and vice versa.

The central bank can thus deter runs by threatening high inflation when nominal spending is excessive—i.e., by limiting the supply of goods in the case of a run, making the run suboptimal.[30] This creates a CBDC trilemma, where the

[29] This is taken as an equivalence result where, in the absence of a banking panic, allocations met with private financial intermediation will also be met with a CBDC.

[30] This is true for the patient depositors, who would not receive utility from consuming goods in the present. However, the inflation threat must be credible for runs not to occur. Note also that high inflation only occurs off the equilibrium path, in contrast to the results of the DD model where runs can occur in equilibrium.

central bank can attain at most two out of the following three goals: efficiency (optimal ex ante risk sharing in the sense of DD), financial stability (the absence of runs), and price stability. If the primary goal is price stability, for example, then either the allocation will be suboptimal or there will always be the risk of destabilizing runs. The trilemma would tend to worsen as well under political-economy pressures. However, observers note that such an extremely centralized economy is unlikely at the present time, with no major central bank considering such features [Auer et al. 2021].

A variety of other models similarly derive from the DD framework, such as those by Skeie [2021] and Popescu [2022], offering insight into the possible effects of central bank issuance of their own digital currencies. The first argues that appropriate and dynamic management of policy rates paid on bank reserves relative to interest rates paid on CBDC support optimal investment and risk sharing and prevent disintermediation of banks and digital currency runs into CBDC. The second focuses on cross-border CBDCs and explores the implications of having a foreign CBDC that serves as an international safe asset, concluding that the presence of such an entity increases the risk of financial disintermediation of the banking system and financial instability marked by high and volatile capital outflows. The findings suggest the importance of coordination in the design of CBDCs at the global level. Other models have investigated the impact of CBDCs on financial stability without using the framework of DD but nevertheless acknowledge their pioneering framework (e.g., Bitter [2020]).

5. Concluding remarks

Evidently, the DD model has been a valuable theoretical contribution, with far-reaching intellectual and policy influence. It has become a solid building block for models featuring different settings and scenarios, yielding important insights for policymakers. Built in the early 1980s, it continues to have a profound impact on today's thinking, even as the financial system has evolved, with technological change introducing new instruments and actors.

Applying the DD framework to emerging markets and developing economies has helped drive home the necessity of certain policies and reforms to lower vulnerability to financial crises. These include broad strokes such as maintaining sound macroeconomic fundamentals (e.g., avoiding unhealthy booms in lending and building up foreign reserves); better handling of capital flows; and proper sequencing of capital liberalization. The expanded framework also points to the need to avoid specific triggers in developing economies—such as unhedged short-term foreign debt, rickety exchange rate pegs, and undercapitalized banks—which could increase the likelihood of coordination failures among investors and susceptibility to short-term debt runs.

Similarly applying DD's framework to modeling or even just examining the evolving financial system allows better analysis of the impact of such changes.

For now, it appears that the rapid rise in fintech may not entail too much risk, though there are worrisome areas, such as the possible re-emergence of shadow banks, which must be addressed. Formal models based on the DD model that seek to understand new financial concepts, such as CBDCs, provide a way to reveal possible blind spots, such as those that could ultimately work to undermine central bank independence.

Built on solid microeconomic foundations, the DD model will likely continue to be applicable despite a rapidly changing financial intermediation landscape. It has captured the key mechanisms in a form that is easy to incorporate in other models as well as to communicate to the layman. Thus, DD's ideas will likely remain an important component of future models. While the risk of a financial crisis will never go away, there is greater confidence about the future now that, with DD's solid research and others that followed, we have a better understanding of financial crises and how to prevent or handle them.

References

Allen, F. and D. Gale [1998] "Optimal financial crises", *Journal of Finance* 53(4):1245-1284.

Allen, F. and D. Gale [2004a] "Financial fragility, liquidity, and asset prices", *Journal of the European Economic Association* 2(6):1015-1048.

Allen, F. and Gale [2004b] "Financial intermediaries and markets", *Econometrica* 72(4):1023-1061.

Anginer, D. and A. Demirguc-Kunt [2018] "Bank runs and moral hazard", *World Bank Policy Research Working Paper No.* 8589.

Auer, R., J. Frost, L. Gambacorta, C. Monnet, and T.R. Shin [2021] "Central bank digital currencies: motives, economic implications and the research frontier", *BIS Working Paper No*. 976. Bank for International Settlements.

Bank of Japan [2023] "The Bank of Japan's policy measures during the financial crisis", https://www.boj.or.jp/en/mopo/outline/cfc.htm. Accessed November 15, 2023.

Basel Committee on Banking Supervision (BCBS) [2018] "Sound practices: implications of fintech developments for banks and bank supervisors", *Bank for International Settlements*. https://www.bis.org/bcbs/publ/d431.pdf.

Bernanke, B.S. [1983] "Nonmonetary effects of the financial crisis in the propagation of the Great Depression", *American Economic Review* 73(3):257-276.

Bernanke, B.S. [2009] "Reflections on a year of crisis", *Board of Governors of the Federal Reserve System*. https://www.federalreserve.gov/newsevents/speech/bernanke20090821a.htm.

Bernanke, B.S. [2018] "The real effects of disrupted credit: evidence from the global financial crisis", *Brookings Papers on Economic Activity* 49(2):251-342.

Bitter, L. [2020] "Banking crises under a central bank digital currency", *ZBW - Leibniz Information Centre for Economics*. https://www.econstor.eu/bitstream/10419/224600/1/vfs-2020-pid-40056.pdf.

Bollaert, H., F. Lopez-de-Silanes, and A. Schwienbacher [2021] "Fintech and access to finance", *Journal of Corporate Finance* 68(C):1-14.

Brunnermeier, M. and A. Krishnamurthy [2020] "COVID crisis response: 'evergreening proposal'", *World Bank*. https://thedocs.worldbank.org/en/doc/278551588189773 8460050022020/original/20200429WorldBankCOVIDProposalbyMarkusb.pdf.

Bryant, J. [1980] "A model of reserves, bank runs, and deposit insurance", *Journal of Banking and Finance* 4(4):335-344.

Burnside, C., M. Eichenbaum, and S. Rebelo [2004] "Government guarantees and self-fulfilling speculative attacks", *Journal of Economic Theory* 119(1):31-63.

Calomiris, C.W. and G. Gorton [1991] "The origins of banking panics: models, facts, and bank regulation", in R.G. Hubbard (ed.), *Financial markets and financial crises* (109-174). National Bureau of Economic Research.

Calvo, G. [2012] "Financial crises and liquidity shocks a bank-run perspective", *European Economic Review* 56(3):317-326.

Carlsson, H. and E. Van Damme [1993] "Global games and equilibrium selection", *Econometrica* 61(5):989-1018.

Chang, R. and A. Velasco [1998] "Financial crises in emerging markets: a canonical model", *Federal Reserve Bank of Atlanta Working Paper No*. 98-10:1-50.

Chang, R. and A. Velasco [2000] "Liquidity crises in emerging markets: theory and policy", in B.S. Bernanke and J.J. Rotemberg (eds.), *NBER macroeconomics annual 1999* (Vol. 14, 11-78). MIT Press.

Chang, R. and A. Velasco [2001] "A model of financial crises in emerging markets", *Quarterly Journal of Economics* 116(2):489-517.

Committee for the Prize in Economic Sciences in Memory of Alfred Nobel [2022] "Scientific background on the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2022: financial intermediation and the economy", The Royal Swedish Academy of Sciences.

Cooper, R. and T.W. Ross [2002] "Bank runs: deposit insurance and capital requirements", *International Economic Review* 43(1):55-72.

Diamond, D.W. [1984] "Financial intermediation and delegated monitoring", *Review of Economic Studies* 51(3):393-414.

Diamond, D.W. [1997] "Liquidity, banks, and markets", *Journal of Political Economy* 105(5):928-956.

Diamond, D.W. [2022] "Financial intermediation and financial crises (Nobel prize lecture)", *The Royal Swedish Academy of Sciences*. https://www.nobelprize.org/uploads/2022/11/diamond-lecture-slides.pdf.

Diamond, D.W. and P.H. Dybvig [1983] "Bank runs, deposit insurance, and liquidity", *Journal of Political Economy* 91(3):401–419.

Diamond, D.W. and P.H. Dybvig [1986] "Banking theory, deposit insurance, and bank regulation", *Journal of Business* 59(1):55-68.

Didier, T., F. Huneeus, M. Larrain, and S.L. Schmukler [2021] "Financing firms in hibernation during the COVID-19 pandemic", *Journal of Financial Stability* 53:1-14.

Dooley, M.P. [2000] "A model of crises in emerging markets", *Economic Journal* 110:256-272.

Eichengreen, B. and A.K. Rose [1998] "Staying afloat when the wind shifts: external factors and emerging-market banking crises", *NBER Working Paper Series No.* 6370.

Farhi, E., M. Golosov, and A. Tsyvinski [2009] "A theory of liquidity and regulation of financial intermediation", *Review of Economic Studies* 76(3):973-992.

Fernández-Villaverde, J., D. Sanches, L. Schilling, and H. Uhlig [2021] "Central bank digital currency: central banking for all?", *Review of Economic Dynamics* 41:225-242.

Frankel, J.A. [2010] "Monetary policy in emerging markets: a survey", *NBER Working Paper No*. 16125.

Furman, J. and J.E.Stiglitz [1998] "Economic crises: evidence and insights from East Asia", *Brookings Papers on Economic Activity* 1998(2):1-135.

Gambacorta, L. [2023] *Digital banking: opportunities and challenges*. Bank for International Settlements.

Goldstein, I. and A. Pauzner [2005] "Demand–deposit contracts and the probability of bank runs", *Journal of Finance* 60(3):1293-1327.

Gorton, G. [1988] "Banking panics and business cycles", *Oxford Economic Papers* 40(4):751-781.

Hazlett, D. [1997] "Deposit insurance and regulation in a Diamond-Dybvig banking model with a risky technology", *Economic Theory* 9(3):453-470.

He, Z. and Y. Ma [2022] "The insight and influence of Douglas Diamond", *Promarket*. https://www.promarket.org/2022/10/11/the-insight-and-influence-of-douglas-diamond/.

Jacklin, C.J. [1987] "Demand deposits, trading restrictions, and risk sharing", in E.C. Prescott and N. Wallace (eds.), *Contractual arrangements for intertemporal trade* (pp. 26-47). University of Minnesota Press.

Jeanne, O. [2009] "Debt maturity and the international financial architecture", *American Economic Review* 99(5):2135-2148.

Jeasakul, P., C.H. Lim, and E. Lundback [2014] "Why was Asia resilient? Lessons from the past and for the future", *IMF Working Paper No*. WP/14/38.

Kashyap, A.K. [2015] "Bank runs aren't madness: this model explained why", *Chicago Booth Review*. https://www.chicagobooth.edu/review/bank-runs-arent-madness-this-model-explained-why.

Krugman, P. [1999] "Balance sheets, the transfer problem, and financial crises", *International Tax and Public Finance* 6:459-472.

Kuroda, H. [2020] "Japan's economy and monetary policy: speech at a meeting with business leaders in Osaka", *Bank of Japan*. https://www.boj.or.jp/en/about/press/koen_2020/ko200923a.htm.

Masson, P.R. [1999] "Multiple equilibria, contagion, and the emerging market crisis", *IMF Working Paper No*. 99/164.

Matsuoka, T. [2018] "Banks and liquidity crises in emerging market economies", *Journal of Economic Dynamics & Control* 94(C):43-62.

McKinnon, R.I. and H. Pill [1997] "Credible economic liberalizations and overborrowing", *American Economic Review* 87(2):189-193.

Morris, S. and H.S. Shin [2000] "Rethinking multiple equilibria in macroeconomic modeling", University of Chicago Press on behalf of the The National Bureau of Economic Research.

Murinde, V., E. Rizopoulos, and M. Zachariadis [2022] "The impact of the FinTech revolution on the future of banking: opportunities and risks", *International Review of Financial Analysis* 81(11):1-27.

Navaretti, G.B., G. Calzolari, J.M. Mansilla-Fernandez, and A.F. Pozzolo [2018] "Fintech and banking: friends or foes?", https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3099337.

Noman, A.H., M.K. Hassan, S. Pervin, C.R. Isa, and C. Sok-gee [2022] "The mediating role of competition on deposit insurance and the risk-taking of banks in ASEAN countries", *Research in International Business and Finance* 59(C):1-19.

Popescu, A. [2022] "Cross-border central bank digital currencies, bank runs and capital flows volatility", *IMF Working Paper No*. WP/22/83.

Prescott, E.S. [2010] "Introduction to the special issue on the Diamond-Dybvig model", *Economic Quarterly* 96(1):1-9.

Radelet, S. and J.D. Sachs [1998] "The East Asian financial crisis: diagnosis, remedies, prospects", *Brookings Papers on Economic Activity* 1998(1):1-90.

Rodrik, D. and A. Velasco [1999] "Short-term capital flows", *NBER Working Paper No*. 7364.

Sachs, J. [1996] "Alternative approaches to financial crises in emerging markets", *Brazilian Journal of Political Economy* 16(1):202-216.

Samuelson, P. [1958] "An exact consumption-loan model of interest with or without the social contrivance of money", *Journal of Political Economy* 66(6):467-482.

Schilling, L., J. Fernández-Villaverde, and H. Uhlig [2020] "Central bank digital currency: when price and bank stability collide", *NBER Working Paper No*. 28237.

Skeie, D.R. [2021] "Digital currency runs", *WBS Finance Group Research Paper*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3294313.

Thakor, A.V. [2020] "Fintech and banking: what do we know?", *Journal of Financial Intermediation* 41(C):1-13.

Walker, M. [2017] "The enduring relevance of Diamond-Dybvig", https://olinblog.wustl.edu/2017/11/everyone-wants-piece-banks-fail/.

Working Group established by the Committee on the Global Financial System and the Financial Stability Board [2017] "Fintech credit: market structure, business models and financial stability implications", *Financial Stability Board*. https://www.bis.org/publ/cgfs_fsb1.htm.

The Philippine Review of Economics
60(2):64-80. DOI:10.37907/4ERP3202D

Toward a general neoclassical theory of economic growth

Delano S. Villanueva*
International Monetary Fund**

The Harrod-Domar (H-D) growth model assumes a fixed capital-output ratio, signifying absence of substitutability between capital and labor, leading to a "knife-edge" problem wherein balanced growth of capital (fixed warranted rate) and labor (fixed natural rate) occurs only by accident, preventing the attainment of macroeconomic stability with full employment. The neoclassical Solow-Swan (S-S) growth model provides an elegant solution to the H-D problem by endogenizing the warranted rate via the saving-investment relation, wherein capital growth is a function of a fully adjusting income-capital ratio (inverse of the H-D capital-output ratio)—allowing for smooth substitutability between capital and labor while keeping the natural rate exogenously fixed. The S-S model implies a positive, albeit *temporary* output growth effect of a higher saving rate. The present paper extends the capital-labor ratio's influence onto the natural rate via effects on labor productivity through a modified Arrow learning by doing framework, and via labor participation through real wage adjustments. Thus, the positive output growth of a higher saving rate, although *temporary* in the short run as in the S-S model, is *permanent* in the long run through adjustments in both the warranted and natural rates—a generalization of the Solow-Swan model.

JEL classification: E130, O410
Keywords: Harrod-Domar, neoclassical growth model, Solow-Swan, warranted rate, natural rate, balanced growth, learning by doing, labor participation

1. Introduction

The basic Solow [1956]-Swan [1956] or S-S growth model provides an elegant solution to the "knife-edge" Harrod [1939]-Domar [1946] or H-D problem[1] by endogenizing the warranted rate via the saving-investment relation, wherein capital growth is a function of a fully adjusting income-capital ratio (inverse of the H-D capital-output ratio)—allowing smooth substitutability between capital and labor while keeping the natural rate exogenously fixed. This makes the warranted

* Address all correspondence to dansvillanueva@gmail.com.
** Former Advisor, IMF
[1] Explained below.

rate a negative[2] function of the capital-labor ratio, allowing short-run adjustments in output or income during the transition to the exogenously fixed natural rate—sum of labor augmenting productivity/technical change and working population growth. The present paper extends the effects of the capital-labor ratio to the natural rate via capital intensity effects on labor augmenting productivity and labor participation. Thus, equilibrium growth is obtained through adjustments in both warranted and natural rates—a generalization of the S-S growth model.

The model is not meant to take account of aggregate demand—neither is the S-S model. Like the S-S model, it is a supply-side model, albeit a generalized one (where the natural rate is made endogenous via endogenous labor productivity and labor participation). Questions on real GDP, inflation, and unemployment, as opposed to *capacity or potential* GDP, are not addressed. For a merger of the textbook short-run macroeconomics of aggregate demand and the S-S textbook long-run macroeconomics of aggregate supply, as called for by Solow [2022] in the 1992 addendum to his Nobel Prize speech , and for a numerical application to the Philippines, see Villanueva et al. [2023], Chapters 7 and 10, respectively.

To put this paper in the simplest perspective, assume a constant-returns (unit-homogeneous) aggregate production function,

$$Y = K^{\alpha}L^{(1-\alpha)}, \tag{1}$$

where Y is output, K is capital stock, $L = APN$ is effective labor (in efficiency units), A = a labor productivity or technology multiplier, P = labor participation rate, N = total population, α = elasticity of output with respect to capital, $1-\alpha$ = output elasticity with respect to labor, and $0 < \alpha < 1$ is a constant. Conlisk [1968] had shown that in a general production function $Y = F(K, L)$ with constant rates of factor-augmenting technical change attached to K and L and subject to degree β returns to scale,[3] the existence of a well-behaved and balanced growth equilibrium[4] requires a unitary elasticity of substitution $\varepsilon(k) = (\dot{k}/k)/(\dot{u}/u)$ where $u = F_2/F_1$ and $k = K/L$.[5] If F is Cobb-Douglas ($\beta = 1$) as in (1), then $\varepsilon(k) = 1$ and α is the constant income share of capital.[6]

Income growth at any moment of time is equal to the weighted sum of capital growth and labor growth, the weights being α and $1-\alpha$, respectively (a dot over a variable denotes time derivative, i.*e*., $\dot{K} = dK/dt$):

$$\frac{\dot{Y}}{Y} = \alpha\frac{\dot{K}}{K} + (1-\alpha)\frac{\dot{L}}{L}, \tag{2}$$

[2] Owing to diminishing marginal product of capital. The H-D growth model assumes a fixed capital-output ratio, signifying absence of substitutability between capital and labor.

[3] $\beta < 0$ signifies decreasing returns to scale, $\beta = 1$ constant returns to scale, and $\beta > 1$ increasing returns to scale.

[4] Defined as g^* = where $(\dot{K}/K)^* = (\dot{L}/L)^* = (\dot{Y}/Y)^* = (\dot{A}/A)^* + (\dot{P}/P)^* + (\dot{N}/N)^*$, " * " indicates equilibrium and $(\dot{K}/K)^*$, $(\dot{A}/A)^*$, $(\dot{P}/P)^*$, and $(\dot{Y}/Y)^*$ are functions of capital intensity $k^* = K/L$, as postulated in Section 2.

[5] This requirement applies to growth models with *increasing returns to capital* ($\beta > 1$), e.g., Romer [1986], Lucas [1988], Grossman and Helpman [1990, 1991], Rivera-Batiz and Romer [1991], Barro and Sala-i-Martin [1995], and Aghion and Howitt [1998]. See discussion in Conlisk [1968].

[6] See Chapter 6 of Villanueva et al. [2023] and Conlisk [1968].

Capital growth is the *warranted rate* and labor growth is the *natural rate*. The warranted rate (saving-investment) is a monotonically declining function of the capital-labor ratio,

$$\frac{\dot{K}}{K} = s\,\frac{Y}{K} - \delta, \tag{3}$$

where s is the fixed gross saving/income ratio and δ is a constant depreciation rate.[7] From the definition $L = APN$, labor growth or the natural rate, is given by

$$\frac{\dot{L}}{L} = \frac{\dot{A}}{A} + \frac{\dot{P}}{P} + \frac{\dot{N}}{N}\,. \tag{4}$$

Assume, as in the S-S model, that $\dot{A}/A = \lambda$, $\dot{P}/P = 0$, and $\dot{N}/N = n$, where λ and n are constants. In the steady state, k is constant at k^*,[8] implying that

$$\frac{\dot{K}}{K}^* = s\,\frac{Y}{K}^* - \delta = \frac{\dot{L}}{L}^* = \lambda + n \tag{5}$$

and by the constant-returns assumption, using (1) and definition, $g^* = \lambda + n$,

$$\frac{\dot{K}}{K}^* = sk^{*(\alpha-1)} - \delta = \frac{\dot{L}}{L}^* = \frac{\dot{Y}}{Y}^* = g^* = \lambda + n, \tag{6}$$

where g^* defines the equilibrium growth rate of income Y, or the balanced growth path.

The knife-edge H-D problem ($\dot{P}/P = 0$ by assumption) is expressed by the condition

$$\frac{s}{v} - \delta \gtreqless \lambda + n, \tag{7}$$

where $v = K/Y$ is the fixed H-D capital-output ratio in (5). Since both sides are constants, equilibrium growth is not assured.[9]

The S-S model solves the knife-edge H-D problem by employing a neoclassical production function with smooth substitutability between capital and labor, i.e., $1/v$ is a monotonically decreasing function of k, such that the warranted rate fully adjusts from any initial level of k, making balanced growth possible.[10] However, equilibrium income growth g^* remains exogenous because the natural rate, being fixed at $\lambda + n$, serves as a bottleneck to the growth process, making the positive growth effect of a higher saving rate s temporary.[11]

The present paper extends the capital-labor ratio influence on the natural rate via capital intensity effects on labor productivity ($\dot{A}/A$) through a modified

[7] The income-capital ratio Y/K declines with k owing to diminishing marginal product of capital—(3).
[8] The Inada [1963] conditions assure the existence of a unique value of k^*.
[9] Equality signifies simultaneous achievement of balanced growth, macroeconomic stability, and full employment. Inequality signifies either inflationary spiral or continuous deflation with unemployment.
[10] Refer to footnote 7.
[11] As Solow [1991:4] calls it.

Arrow [1962] learning by doing framework [Villanueva 1994], and on labor participation ($\dot{P}/P$) through real wage [Villanueva 2020]. Thus, equilibrium income growth is obtained through adjustments in both the warranted rate and in the natural rate.

The 1960s through 1990s saw attempts to solve the S-S model's exogeneity of the natural rate $\dot{A}/A + \dot{P}/P + \dot{N}/N$, by making labor-augmenting technical change $\dot{A}/A$ endogenous.[12] An early learning by doing model by Arrow [1962] assumes that learning has a positive effect on the equilibrium growth of output. If labor productivity A changes according to $\dot{A}/A = \varnothing(\dot{K}/K)$, where $\varnothing$ is a learning coefficient, then equilibrium output growth g^* is a multiple of the S-S steady-state growth rate $n + \lambda$, since $0 < \varnothing < 1$ by assumption:

$$\frac{\dot{K}}{K}^* = \frac{\dot{L}}{L}^* = \frac{\dot{Y}}{Y}^* = g^* = (n + \lambda) / (1 - \varnothing). \tag{8}$$

Note the absence of the saving rate s in (8). However, if the Arrow [1962] model is interpreted as a change in learning $\dot{A}/A$ being proportional to the capital-labor ratio $k = K/L$,[13] and not to the growth rate of the capital stock $\dot{K}/K$, then the present paper is the Arrow model extended to the case of endogenous labor participation $\dot{P}/P$ (elaborated below).

Nelson and Phelps [1966], Conlisk [1967], and Villanueva [1994] advanced early growth models with endogenous labor-augmenting technical change, deriving the key result that an increase in the saving rate raises equilibrium output growth. Agénor [2004:466-471] refers to Villanueva's [1994] model as a variant of the Conlisk [1967] model and "an extension of Arrow's [1962] learning by doing model,…[wherein] the productivity of workers increases when the relative availability of capital goods (for instance, the stock of high-performance computers) rises", leading to enhanced equilibrium growth effects of saving and investment rates. More precisely, Villanueva [1994] interprets Arrow's [1962] learning by doing model by positing $\dot{A}/A = \theta k + \lambda$, where $\theta > 0$ is a learning coefficient and λ is an exogenous labor-augmenting productivity/technical change term.

Subsequent contributions constructed increasingly complex models. Romer [1986; 1990] posited a knowledge-producing sector, alongside a goods-producing sector. The stock of knowledge is a non-rival good—its use in one sector does not preclude its use in the other sector. Lucas [1988] proposed models emphasizing human capital accumulation through schooling and learning by doing, but he abstracted from the economics of demography.[14] Grossman and Helpman [1991] focused on innovation financed by investments in industrial research. Rebelo's [1991] AK model assumed that all productive inputs, including human capital,

[12] For an engaging history of endogenous growth theory, see Warsh [2007]. Solow [1991] has been critical of endogenous growth models with their emphasis on endogenous technical change and increasing returns. See Chapter 1 of Villanueva et al. [2023] and the third paragraph of the present introductory section.

[13] This is my interpretation of the Arrow [1962] model.

[14] Lucas [1988:6] admits that this is a serious omission.

are reproducible.[15] Aghion and Howitt [1998] highlighted imperfect markets in the research and development (R&D) sector and Schumpeterian creative destruction. The knowledge-innovation-R&D sector is assumed to be subject to increasing returns so that growth does not vanish. Conlisk [1967] had shown that increasing returns to capital yield explosive growth, which rarely or temporarily happens in the real world.[16] He had demonstrated that a growth model with endogenous labor-augmenting technical change and an aggregate production function that is subject to diminishing returns to capital is consistent with positive and permanent growth effects of an increased saving rate (or of any change in the other model parameters with expected signs).

In all the above growth models, the labor participation rate P is an exogenous constant fraction or percentage by assumption.[17] Another solution to the knife-edge problem, besides the S-S model's variable capital-output ratio implicit in a well-behaved neoclassical production function with smooth factor substitution and wage-rice flexibility and endogenous labor-augmenting productivity multiplier [Villanueva 1994], is a fully-adjusting natural rate via an endogenously-determined labor participation rate P [Villanueva 2020]. In a carefully researched IMF empirical study, Grigoli et al. [2018:18] found robust results indicating that, among others, an increasingly educated[18] labor force influenced significantly and positively the labor participation rates in 36 advanced economies. Referring to the US in particular, the Congressional Budget Office (CBO) [2018] issued a working paper on labor participation, containing similar explanatory variables included in the IMF study, and arriving at similar statistical results. The study noted that the US labor participation rate began an uninterrupted decline in the latter half of the 1990s, coinciding with the aging and retirement of baby boomers. In 2007, the labor participation rate stood at 66 percent. A decade later, in 2017 Q4, it fell to 63.2 percent. The CBO projects that the US labor participation rate will continue to decline and will be 60.2 percent in 2028 Q4. The projected increase in educational attainment, which has a positive effect on the labor participation rate, will not be enough to offset the continued decline attributed to aging and retirements and to the stagnation in real wages, among other factors.

Motivated by the empirical findings of Grigoli et al. [2018] and the CBO [2018], Villanueva [2020] postulated that the proportionate change in labor

[15] Output $Y = AK$, where Y is constant returns to capital K, implying that Y always grows at the same rate as K, and is equal to s^*A, where s^* is the fraction of income saved for investment in physical *and* human capital ($s^* > s$; s is income saved for investment in physical capital) and A is a technological constant. This property is in sharp contrast to the transitional growth dynamics in the S-S model.

[16] However, see Conlisk [1968] and the discussion in the third paragraph of the present introductory section.

[17] Whether it is 70 percent or any other percentage, the rate of change in P is assumed to be zero. The labor participation rate and unemployment rate are metrics used to gauge the health of the labor market. The key difference between the two indicators is that the participation rate measures the percentage of people who are in the labor force, while the unemployment rate measures the percentage within the labor force that is currently unemployed.

[18] Workers with secondary and tertiary degrees.

participation consists of exogenous components including aging and retirements and endogenous components including aggregate income per man-hour and real wages. The objective was to generalize the equilibrium property of the S-S model by making the natural rate fully adjusting through endogenous labor participation [Villanueva 2020], additional to endogenous learning by doing that improves labor productivity [Villanueva 1994].

The rest of the paper is organized as follows. Section 2 presents and discusses the general neoclassical growth model, followed by the analytics of the temporary and permanent growth effects of changes in the structural parameters, notably, the saving rate, learning coefficient, labor-augmenting productivity/technical change, and components of labor participation. Section 3 concludes.

2. A general neoclassical growth model

Equations (9)-(16) below comprise the general neoclassical growth or extended (*e*) model:

$$Y = K^{\alpha} L^{(1-\alpha)} \tag{9}$$

$$L = APN \tag{10}$$

$$\frac{\dot{K}}{K} = s\frac{Y}{K} - \delta \tag{11}$$

$$\frac{\dot{A}}{A} = \theta k + \lambda \tag{12}$$

$$\frac{\dot{P}}{P} = \beta + \rho\left(\frac{Y}{L}\right) + \omega RW \tag{13}$$

$$\frac{\dot{N}}{N} = n \tag{14}$$

$$RW = \frac{\partial Y}{\partial L} \tag{15}$$

$$k = \frac{K}{L} \tag{16}$$

where $Y = GDP$; K = capital; L = effective labor; A = Harrod-neutral productivity or technical change multiplier; P = labor participation rate; N = population base; RW = real wage rate; k = capital/labor ratio ; α = output elasticity with respect to capital; (1-α) = output elasticity with respect to labor; s = gross saving rate; δ = depreciation rate; θ = learning coefficient; λ = constant rate of exogenous Harrod-neutral technical change; β = exogenous, noneconomic determinants of labor participation; and ρ, ω, n = constant parameters.

Equation (9) repeats (1) with the same properties. Equation (10) defines effective labor L as the product term APN. Equations (11)-(14) are the dynamic relationships governing rates of change in K, A, P, and N. Equation (15) is a profit maximization condition that the real wage be set equal to labor's marginal product. Finally, (16) defines capital intensity as the ratio of K to L.

Equation (11), the warranted rate, repeats Equation (3). The derivation of the natural rate $\dot{L}/L$ is as follows: Villanueva [1994] interprets Arrow's [1962] learning by doing model by positing $\dot{A}/A = \theta k + \lambda$, as in (12), where $\theta > 0$ is a learning coefficient. The idea is that as the per capita stock of capital K/N with embodied advanced technology gets larger, the learning experience makes workers more productive.[19]

Reflecting the empirical findings of Grigoli et al. [2018] and the CBO [2018], (13) states that the proportionate change in labor participation P is the sum of an exogenous component β and endogenous components $\rho(Y/L)$ and ωRW, while (14) expresses the standard rate of exogenous population growth. The exogenous term β includes aging and retirements; changes in labor market policies and institutions, e.g., tax-benefits (tax credits and unemployment benefits); and a host of non-economic variables identified in the aforementioned empirical studies.[20] The endogenous terms are: (a) the portion of aggregate income per man-hour (Y/L) spent on secondary and tertiary education and its effect on the number of graduates, and the latter's effect on the labor participation rate;[21] and (b) the real wage RW that, under profit maximization, is equal to labor's marginal product $\partial Y/\partial L = (1-\alpha)k^{\alpha}$. Equation (16) defines k as the capital/labor ratio. Increases in the percentage ρ of aggregate income per man-hour (Y/L) spent on secondary and tertiary education and in the real wage ($\omega > 0$) are expected to raise the rate of labor participation. There are eight equations with eight endogenous variables—Y, K, L, A, P, N, RW, and k (time t is suppressed).

2.1. Reduced model

Using (9) and (16), (11) becomes

$$\frac{\dot{K}}{K} = sk^{(\alpha-1)} - \delta. \tag{17}$$

Differentiating (9) with respect to L and substituting (15)-(16) yield

$$RW = (1-\alpha)k^{\alpha}. \tag{18}$$

[19] Using $L = AN$, letting $P = 1$, rewrite the above equation as $\dot{A} = \theta(K/N) + \lambda A$, or $(\dot{A}/A) = \theta k + \lambda$, where $k = K/L$.

[20] If β denotes aging and retirements, then the growth effect is negative; if β denotes tax credits and unemployment benefits, then the growth effect is positive.

[21] The coefficient $\rho > 0$ is a composite parameter reflecting the fraction of aggregate income spent on secondary and tertiary education and its effect on the number of graduates, and the latter's effect on labor participation.

Using (9) and (16) yields

$$\frac{Y}{L} = k^{\alpha\text{-}1}. \tag{19}$$

Equations (18)-(19) into (13) yield

$$\frac{\dot{P}}{P} = \beta + [\rho + \omega\ (1\text{-}\alpha)]k^{\alpha}. \tag{20}$$

Time differentiating (10), using (12), (14), and (20), yields

$$\frac{\dot{L}}{L} = \theta k + [\rho + \omega\ (1\text{-}\alpha)]k^{\alpha} + \lambda + n + \beta. \tag{21}$$

The equilibrium growth rate of per capita income is

$$g^{*} - n = \theta k_{e}^{*} + [\rho + \omega\ (1\text{-}\alpha)]k_{e}^{*\alpha} + \lambda + \beta, \tag{22}$$

where k^{*} = equilibrium capital intensity, and e refers to the extended model.

Time differentiating (16) and substituting (17) and (21) into the result yield the rate of change of capital intensity at any time,

$$\frac{\dot{k}}{k}(e) = sk^{(\alpha\text{-}1)} - \theta k + [\rho + \omega\ (1\text{-}\alpha)]k^{\alpha} + (\lambda + n + \beta + \delta). \tag{23}$$

Time differentiating (19) and substituting (22) into the result yield the growth rate of per capita income at any moment of time,

$$\frac{\dot{L}}{L} - n = g - n = \theta k_{e}^{*} + [\rho + \omega\ (1\text{-}\alpha)]k_{e}^{*\alpha} + \lambda + \beta + \alpha\frac{\dot{k}}{k}(e). \tag{24}$$

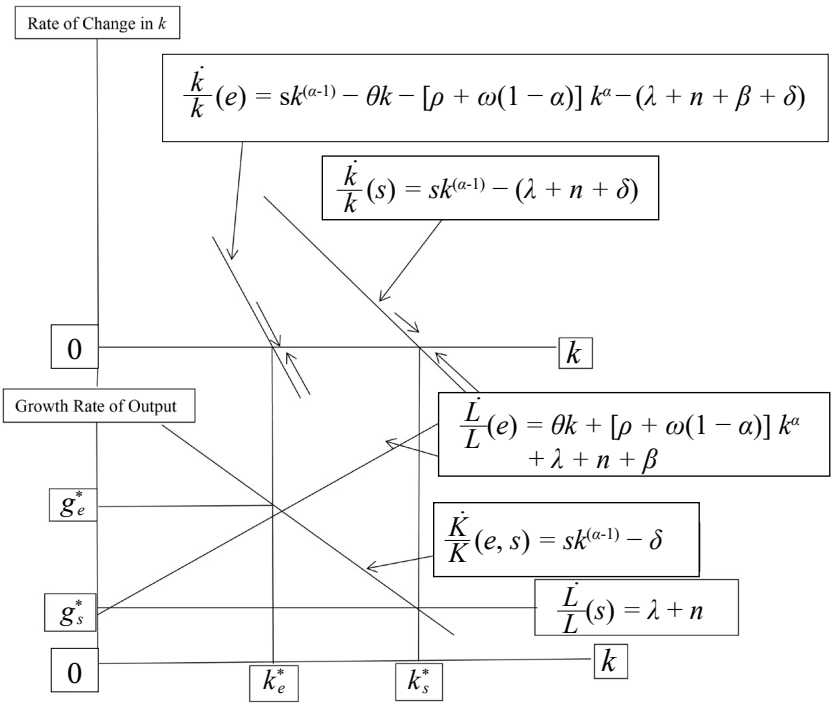
where $\dot{k}/k(e)$ is given by (23). In equilibrium, $\dot{k}/k(e) = 0$ and (24) reduces to (22).

The reduced models in $(\dot{k}/k)$, $(\dot{K}/K)$, and $(\dot{L}/L)$ in the S-S (denoted by s) and extended (denoted by e) models are shown in Figure 1.[22] The upper part shows the proportionate rate of change in the capital-labor ratio and the lower part shows the growth rate of output. In both parts, the horizontal axis shows the level of capital intensity. Given the Inada [1963] conditions, the $\dot{k}/k$ line in either model is downward sloping and intersects the k-axis at some positive k, such as k_{s}^{*} or k_{e}^{*}.When the $\dot{k}/k$ line intersects the k-axis, the equilibrium capital intensity is k_{s}^{*} in the S-S model and k_{e}^{*} in the extended model. In either model, for $k < k^{*}$, $\dot{k}/k > 0$, k increases until it reaches k^{*} at which it becomes constant. For $k > k^{*}$, $\dot{k}/k < 0$, and k decreases until it goes back to k^{*} at which it becomes constant. As capital intensity changes, diminishing marginal and average productivity of capital (in either model) and, in the extended model, positive dependence of labor

[22] Figures 1 to 4 are phase diagrams. Phase diagramming is a powerful tool in analyzing growth models not explicitly involving time.

productivity on learning by doing (a positive function of capital intensity) and of labor participation on capital intensity, provide the economic rationale behind the proportionate changes in capital intensity and in the warranted and natural rates. Specifically, with reference to the lower part of Figure 1, the downward sloping warranted rate line in either model owes to diminishing marginal and average capital productivities as k increases. The horizontal natural rate line in the S-S model ($\dot{L}/L(s)$) reflects the full exogeneity of labor-augmenting productivity. The upward sloping natural rate line in the extended model ($\dot{L}/L(e)$) reflects the positive dependence on capital intensity of efficient labor growth via learning by doing and of labor participation via the real wage (a positive function of capital intensity). As capital intensity rises, more intensive learning leads to greater labor productivity. Higher aggregate income per man-hour translates into higher spending on secondary and tertiary education, higher number of graduates, and higher labor participation. As capital intensity rises, labor's marginal product (real wage) goes up, encouraging higher labor participation and, hence, a larger natural rate. In the lower part of Figure 1, equilibrium growth rates of output g_s^* and g_e^* respectively, in the S-S and extended models, are indicated when the warranted and natural rates are equal at points k_s^* and k_e^*. Note that $g_e^* > g_s^*$ because of endogenous learning by doing and endogenous labor participation in the extended model.

FIGURE 1. Extended (e) and Solow-Swan (s) models



2.2. Temporary and permanent growth

Table 1 shows the permanent effects of changes in the S-S and extended model's parameters on capital intensity and growth rate of income. In the S-S model, only the rates of exogenous labor-augmenting change and population growth have permanent growth effects. In the extended model, higher values for the saving rate, learning coefficient, expenditures on secondary and tertiary education, real wage, exogenous labor-augmenting productivity/technical change, and population growth have growth effects, while aging/retirements and physical capital depreciation impact negatively on growth. On the balanced growth path of the extended model, a constant equilibrium capital/labor ratio means that the warranted rate is equal to the natural rate, and by the constant returns assumption, to the growth rate of per capita output/income as well,

$$\frac{\dot{K}}{K}^{*} = \frac{\dot{L}}{L}^{*} = \frac{\dot{Y}}{Y}^{*} = g^{*} - n = \theta k_e^{*} + [\rho + \omega\,(1\text{-}\alpha)]k_e^{*\alpha} + \lambda + \beta,$$

which is (22). At any moment of time, the output growth rate is a weighted average of the warranted and the natural rates,

$$\frac{\dot{Y}}{Y} = \alpha(sk^{(\alpha\text{-}1)} - \delta) + (1 - \alpha)\{\theta k + [\rho + \omega(1\text{-}\alpha)]\, k^{\alpha} + \lambda + n + \beta\}.$$

There is a divergence between the warranted and natural rates in the short-run transition to equilibrium. In the S-S model, in equilibrium, the natural rate is equal to a constant term $\lambda + n$. In the transition to equilibrium, output growth adjustment falls only on the warranted rate as capital intensity adjusts to its equilibrium value. In the extended model, the equilibrium natural rate adjusts as well to a moving capital/labor ratio.

TABLE 1. Permanent growth

		s	*θ*	*ρ*	*ω*	*β*	*λ*	*n*	*δ*
*k**	S-S	+	0	0	0	0	-	-	-
	Extended	+	-	-	-	-	-	-	-
*g**	S-S	0	0	0	0	0	+	+	0
	Extended	+	+	+	+	+	+	+	-

Notes:

a. s = gross saving rate; θ = learning coefficient; ρ = the portion of aggregate income per man-hour (Y/L) spent on secondary and tertiary education and its effect on the number of graduates, and the latter's effect on the labor participation rate; ω = effect of real wage on labor participation; β = effects on labor participation of aging and retirements, changes in labor market policies and institutions, e.g., tax-benefits (tax credits and unemployment benefits), and a host of non-economic variables; λ = effect of rate of exogenous labor-augmenting productivity/technical change; n = rate of exogenous population growth; and δ = rate of depreciation; k^{*} = equilibrium capital intensity; g^{*} = equilibrium output or income growth.

b. The equilibrium capital intensity k^{*} is the root of $\dot{k}/k\ (e) = sk^{*(\alpha\text{-}1)} - \theta k^{*} - [\rho + \omega(1\text{-}\alpha)]\, k^{*\alpha} - (\lambda + n + \beta + \delta) = \psi(k^{*};\, s,\, \theta,\, \rho,\, \omega,\, \beta,\, \lambda,\, n,\, \delta) = 0$.

c. In the above table, for the extended model $\partial k^{*}/\partial s = -\psi'_{s}/\,\psi'_{k^{*}} > 0$, since $\psi'_{s} > 0$, and $\psi'_{k^{*}} < 0$. The equilibrium output growth rate is $g^{*} = sk^{*(\alpha\text{-}1)} - \delta = H(k^{*};\, s,\, \delta)$, or $g^{*} = \theta k^{*} + [\rho + \omega(1\text{-}\alpha)]\, k^{*\alpha} + \lambda + n + \beta = J(k^{*};\theta,\rho,\omega,\lambda,n,\beta)$. Taking partial derivatives, $\partial g^{*}/\partial s = H'_{k^{*}}\, H'_{s} > 0$; $\partial g^{*}/\partial \delta = H'_{k^{*}} - H'_{\theta} = H'_{k^{*}} - 1 < 0$; $\partial g^{*}/\partial \theta = J'_{k^{*}}\, J'_{\theta^{*}} > 0$. The same procedure was used to derive the signs of the other parameters in the above table.

2.2.1. Growth effects of higher saving rate

Figure 2 shows the effects of an increase in the saving rate on equilibrium capital intensity and equilibrium output growth in the S-S and extended models. The starting equilibrium positions are points $B(k^*_s, g^*_s)$ for the S-S model and $A(k^*_e, g^*_e)$ for the extended model. A higher saving rate shifts the warranted rate line to the right in either model. The new equilibrium positions are indicated by point *C* in the S-S model and point *D* in the extended model. In both models, the capital/labor ratio goes up, albeit the new ratio remains lower in the extended model than in the S-S model, owing to positive labor participation in the extended model. The key difference is that the new equilibrium output growth increases in the extended model but remains unchanged in the S-S model.

The short-run dynamics of the S-S model is taken up first, followed by that of the extended model. During the transition between equilibrium points *B* and *C*, the S-S output growth rate is momentarily higher than the natural rate g^*_s at point *E* because of a higher warranted rate.[23] As noted, Figure 2 repeats the lower panel of Figure 1 (see (17) and (21)). The capital/labor ratio begins to rise from k^*_s to $k^{*\prime}_s$, which slows the warranted rate. Since the natural rate is independent of the capital/labor ratio, only the warranted rate adjusts downward along the segment *EC*.[24] Over time, labor becomes a bottleneck, and the output growth rate slows to the constant natural rate $g^*_s = \lambda + n$ at *C*. At this point, the capital/labor ratio stops rising and stabilizes at a new and higher-level $k^{*\prime}_s$. Thus, the output growth rate effect of a higher saving rate is temporary, and a higher equilibrium capital/labor ratio is the only permanent effect.

In the extended model, following the increase in the saving rate, equilibrium shifts from *A* to *D*. At the starting position *A*, capital grows faster than labor (by the segment *AF*), and the capital/labor ratio rises from k^*_e to $k^{*\prime}_e$. The marginal and average products of capital fall, lowering the level of saving per unit of capital, thus slowing the warranted rate downward along the segment *FD*. On the other hand, the natural rate, instead of remaining constant as in the S-S model, rises because of enhanced learning by doing and higher labor participation associated with a rising capital/labor ratio.[25] This process continues until the warranted and natural rates are again equal via a continuous increase in capital intensity at the new long-run equilibrium *D*, at which point the warranted rate would have fallen to the new and higher value of the natural rate, equal to the new and higher equilibrium output growth rate $g^{*\prime}_e (> g^*_e)$. Thus, the higher output growth effect of a larger saving rate is both temporary (like in the S-S model) and permanent (unlike in the S-S model), the latter owing to the existence of endogenous learning by doing and endogenous labor participation, making the natural rate respond positively to an increase in capital intensity.

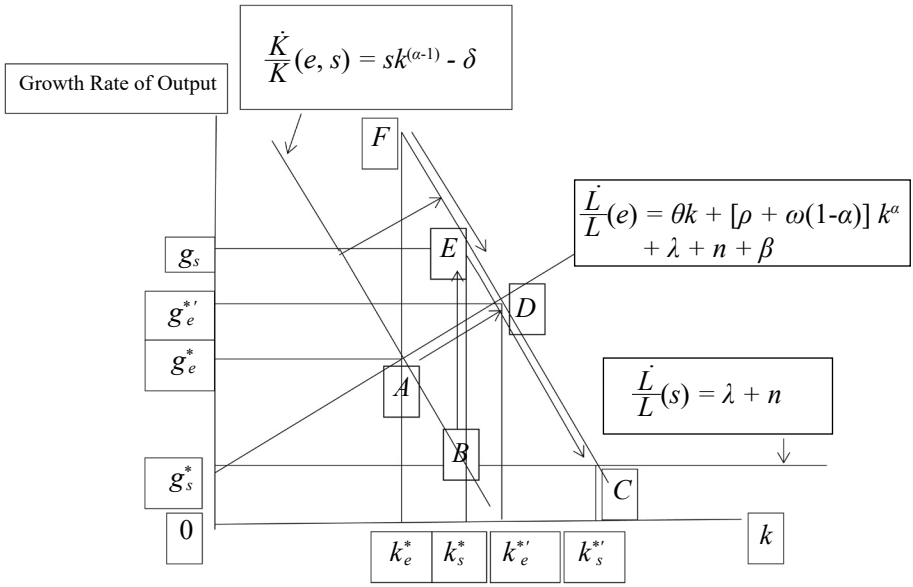
---

[23] The output growth rate at $E = \alpha g_s + (1\text{-}\alpha)g^*_s$.

[24] The output growth rate adjustment is traced by the segment *BEC* in terms of the weighted average of the warranted and natural rates.

[25] The natural rate adjustment is traced by the segment *AD*.

FIGURE 2. Temporary and permanent growth effects of higher saving rate



2.2.2. Growth effects of higher labor-augmenting productivity, enhanced learning by doing

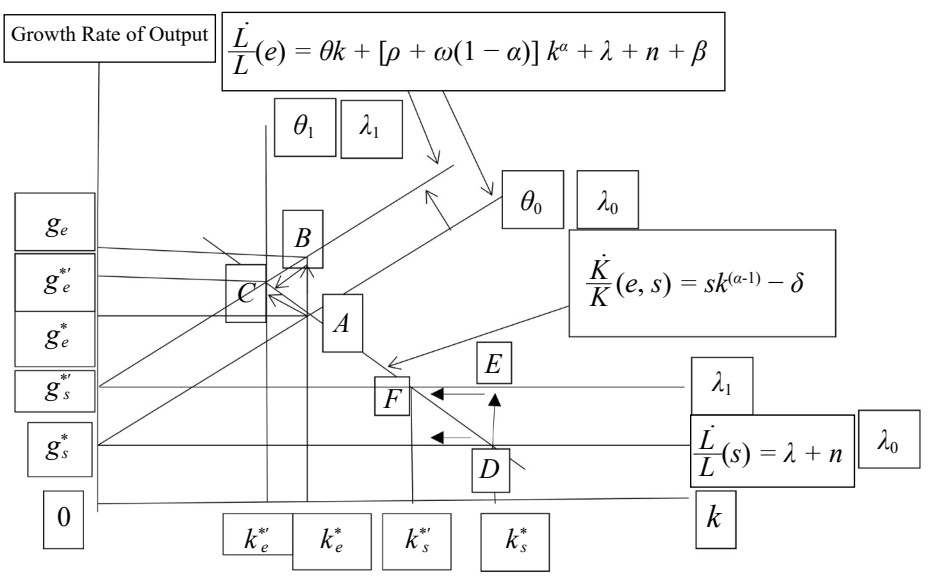
Figure 3 illustrates the temporary and permanent growth effects of enhanced learning by doing and higher exogenous labor-augmenting labor productivity in the S-S and extended models. The growth effects of higher labor-augmenting productivity are taken up first, followed by the growth effects of enhanced learning by doing. The starting equilibrium positions are points $D(\ k_s^*\ ,\ g_s^*\ )$ for the S-S model and $A(\ k_e^*\ ,\ g_e^*\ )$ for the extended model. Higher exogenous labor-augmenting productivity λ shifts the natural rate of the S-S model upward to the $\dot{L}/L = \lambda_1 + n$ line, a parallel shift from the previous line. The extended model's natural rate shifts upward to the left. The new equilibrium positions are indicated by point F in the S-S model and point C in the extended model. In either model, the capital/labor ratio goes down, albeit the new ratio remains lower in the extended model than in the S-S model, owing to positive labor participation and learning by doing in the former. The key difference is that, while the new equilibrium output growth increases to the higher rate of $g_s^{*\prime} = \lambda_1 + n$ in the S-S model, in the extended model the new equilibrium output growth increases to an even higher rate equal to $g_e^{*\prime} = g_s^{*\prime} + \theta k_e^{*} + [\rho + \omega(1-\alpha)]k_e^{*\prime}\ \alpha + \lambda_1 + n + \beta$ (point C).

The short-run (temporary) dynamics of the S-S model is taken up first, followed by that of the extended model. Note that at any moment of time the output growth rate is given by $\dot{Y}/Y = \lambda + n + \alpha\ \dot{k}/k$. Begin with enhanced learning by doing or higher labor-augmenting productivity. Before the steady-state transition between equilibrium points D and F begins, starting from k_s^* the output growth rate jumps

to $\dot{Y}/Y = \alpha g_s^* + (1 - \alpha)\, g_s^{*'} = \lambda_1 + n + \alpha\, \dot{k}/k < g_s^{*'}$ because $\dot{k}/k < 0$ at k_s^*—the natural rate is higher than the warranted rate by segment *ED*. Capital intensity begins to fall from k_s^* to $k_s^{*'}$, resulting in $\dot{k}/k$ turning less and less negative, thus raising output growth until $\dot{k}/k = 0$ at point *F*, wherein the new permanently higher S-S growth rate $\lambda_1 + n$ is reached.

In the extended model, following the increase in λ from λ_0 to λ_1, equilibrium shifts from *A* to *C*. At the starting position A, labor grows faster than capital, and the capital/labor ratio declines from k_e^* toward $k_e^{*'}$. The extended model's output growth rate adjustment is traced by the weighted average of segments *BC* and *AC*, as capital intensity falls from k_e^* to $k_e^{*'}$. The marginal and average products of capital rise, raising the level of saving per unit of capital, accelerating the warranted rate upward along the segment *AC*. On the other hand, the natural rate, instead of remaining constant at $\lambda_1 + n$ as in the S-S model, slows from *B* to *C* because of a lower labor participation rate associated with a declining capital/labor ratio. This process continues until the warranted and natural rates are again equal via a continuous fall in the capital/labor ratio at the new long-run equilibrium *C*, at which point the warranted rate would have risen to the new value of the natural rate, equal to the new and higher equilibrium output growth rate $g_e^{*'}(> g_e^*)$. Thus, whereas in the S-S model the higher output growth is temporary, it is permanent in the extended model because of the existence of endogenous learning by doing and endogenous labor participation.

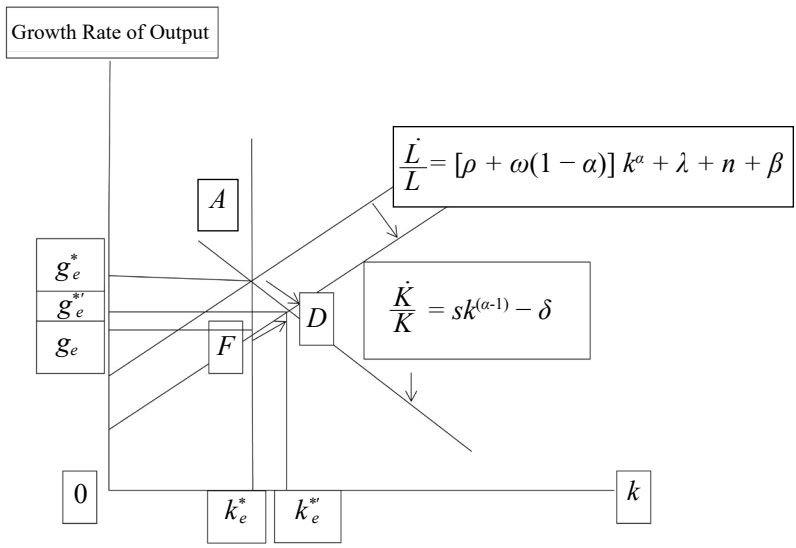
FIGURE 3. Temporary and permanent growth effects of enhanced learning by doing ($\theta_1 > \theta_0$) or higher exogenous labor-augmenting productivity ($\lambda_1 > \lambda_0$)



2.2.3. Growth effects of lower labor participation

Figure 4 illustrates the temporary and permanent growth effects of a decline in labor participation from the CBO [2018] forecast—a lower ρ, ω, or β. The initial equilibrium is at point A, with capital/labor ratio k_e^* and output growth g_e^*. Following the fall in labor participation, the natural rate line shifts downward to the right. Equilibrium shifts from A to D. The capital/labor ratio goes up from k_e^* to $k_e^{*\prime}$, and the equilibrium output growth goes down from g_e^* to $g_e^{*\prime}$. The increase in the equilibrium capital/labor ratio owes to lower effective labor induced by a lower rate of labor participation. The fall in equilibrium output growth is the result of a lower natural rate line along with an unchanged warranted rate line. Notice that there is temporary overshooting of the lower output growth rate at F (in relation to the new permanent growth rate at D). At the starting capital intensity k_e^* the natural rate has a precipitous drop to g_e ($< g_e^{*\prime}$) at F, following the decline in labor participation.[26] As the capital-labor ratio begins to rise from k_e^* to $k_e^{*\prime}$, the natural rate recovers along the segment FD, while the warranted rate falls along the segment AD because of diminishing marginal and average products of capital.

FIGURE 4. Temporary and permanent growth effects of a decline in labor participation (lower *ρ, ω, or β*)



[26] In Figure 4, before k has time to adjust, the output growth rate drops to g_e. Thus, a decline in labor participation results in a short-term contractionary overshooting (a sharp drop of short-run output growth, temporarily even lower than the permanently lower output growth at $g_e^*(< g_e^*)$).

This process continues until the warranted and natural rates are again equal via a continuous increase in the capital/labor ratio at the new long- run equilibrium D, where the permanent output growth $g_e^{*\prime}$ is lower than the initial prevent the decline in, and to encourage higher labor participation through vigorous implementation of public policies on education, on-the-job training, upgrading skills for a digital economy, real wage increases in line with labor productivity, and other labor market participation initiatives identified by Grigoli et al. [2018] and the CBO [2018].

3. Conclusion

This paper's main conclusion is that endogenous learning by doing and endogenous labor participation ensure a fully adjusting natural rate. Together with a fully adjusting warranted rate of the S-S model, the equilibrium growth rate of output and capital intensity are functions of all the structural parameters reflecting saving, learning by doing, and labor participation. There are temporary and permanent growth effects of increases in the saving rate and in the coefficient of learning by doing. A growth-oriented policy includes measures to raise public and private saving rates, as well as expenditures on education and health care aimed at raising labor productivity.

The CBO [2018] projected decline in the labor participation rate over the next decade will result in a permanently lower per capita output growth path, with temporary recessionary overshooting. Policies to restore the previous growth path or to achieve a permanently higher growth path involve avoiding the projected fall in labor participation by aggressive and calibrated spending on secondary and tertiary education, on-the-job training, and skills upgrade to a full-fledged digital economy (a higher ρ), steady increases in real wages in line with labor productivity (a higher ω), vigorous labor market participation activities and more generous tax-benefits (higher β), in order to offset the negative effects of aging and retirements (lower β).

Acknowledgments. I am indebted to my mentor, the late John Conlisk who pioneered endogenous growth theory, and to two anonymous referees for valuable comments. The usual disclaimer applies.

References

Agenor, P. [2004] *The economics of adjustment and growth*. Cambridge, MA: Harvard University Press.

Aghion, P. and P. Howitt [1998] *Endogenous growth theory*. Cambridge, MA: MIT Press.

Arrow, K. [1962] "The economic implications of learning by doing", *Review of Economic Studies* 29:155-173.

Barro, R. and X. Sala-i-Martin [1995] *Economic growth*. NY: McGraw-Hill.

Congressional Budget Office (CBO) [2018] "CBO's projection of labor force participation rates", *Working Paper No*. 2018-04.

Conlisk, J. [1967] "A modified neoclassical growth model with endogenous technical change", *Southern Economic Journal* 34:199-208.

Conlisk, J. [1968] "Non-constant returns to scale in a neoclassical growth model", *International Economic Review* 9:369-373.

Domar, E. [1946] "Capital expansion, rate of growth, and employment", *Econometrica* 14:137-147.

Grigoli, F., Z. Koczan, and P. Tapalova [2018] "Drivers of labor participation: macro and micro evidence", *International Monetary Fund Working Paper* WP/18/150. Washington, DC.

Grossman, G. and E. Helpman [1990] "Comparative advantage and long-run growth", *American Economic Review* 80:796–815.

Grossman, G. and E. Helpman [1991] *Innovation and growth in the global economy*. Cambridge, MA: MIT Press.

Harrod, R. [1939] "An essay in dynamic theory", *The Economic Journal* 49:14-33.

Inada, K-I. [1963] "On a two-sector model of economic growth: comments and generalization", *Review of Economic Studies* 30:119-127.

Lucas, R. [1988] "On the mechanics of economic development", *Journal of Monetary Economics* 22:3-42.

Nelson, R. and E. Phelps [1966] "Investment in humans, technological diffusion, and economic growth", *American Economic Review* 56:69-75.

Rebelo, S. [1991] "Long-run policy analysis and long-run growth", *Journal of Political Economy* 99:500-21.

Rivera-Batiz, L. and R. Romer [1991] "International trade with endogenous technical change", *NBER Working Paper* 3594. Washington: National Bureau of Economic Research.

Romer, P. [1986] "Increasing returns and long-run growth", *Journal of Political Economy* 94:1002-37.

Romer, P. [1990] "Endogenous technical change", *Journal of Political Economy* 98:s71-s102.

Solow, R. [1956] "A contribution to the theory of economic growth", *Quarterly Journal of Economics* 70:65-94.

Solow, R. [1991, November 12] *Policies for economic growth* [Ernest Sturc memorial lecture]. *SAIS*, Johns Hopkins University.

Solow, R. [2022]. *Nobel Prize outreach AB* [Nobel Prize lecture].

Swan, T. [1956] "Economic growth and capital accumulation", *Economic Record* 32:334-362.

Villanueva, D. [1994] "Openness, human development, and fiscal policies: effects on economic growth and speed of adjustment", *IMF Staff Papers* 41:1-29. Reprinted in D. Villanueva et al. [2023] "Chapter 6", *Economic adjustment and growth: theory and practice*. Singapore: World Scientific.

Villanueva, D. [2020] "A modified neoclassical growth model with endogenous labor participation", *Bulletin of Monetary Economics and Banking* 23(1):83-100. Reprinted in D. Villanueva et al. [2023] "Chapter 3", *Economic adjustment and growth: theory and practice*. Singapore: World Scientific.

Villanueva, D., R. Mariano, and C. Guinigundo [2023] *Economic adjustment and growth: theory and practice*. Singapore: World Scientific.

Warsh, D. [2007] *Knowledge and the wealth of nations: a story of economic discovery*. New York: WW Norton and Company.

The Philippine Review of Economics
60(2):81-103. DOI:10.37907/5ERP3202D

Measuring fiscal policy sustainability in developing Asia: what does the Markov Switching Augmented Dickey-Fuller Test tell us?*

Dannah Ysabel M. Premacio**
Ezra Rebecca G. Vidar
Toby C. Monsod
University of the Philippines

This paper measures fiscal sustainability in 22 developing Asian countries for the period 1999–2017. Previous literature generates conflicting results: one paper applies the usual stationarity and cointegration tests and finds that fiscal policy is sustainable but in weak form. Another paper employs a fiscal reaction function and finds that fiscal policy is unsustainable. This paper uses an expanded version of the Markov Switching Augmented Dickey-Fuller test (MS-ADF), which remedies the shortcomings of conventional stationarity tests to provide more statistical power in the presence of nonlinearities and structural breaks. The MS-ADF has never been applied to this set of countries. Results show that the majority of the countries have "uncertain" debt trajectories, not definitively sustainable or unsustainable but somewhere in-between. This is a more nuanced picture of the debt trajectories in the region relative to what is obtained using the established methods. A more nuanced assessment could lead to more suitable policy corrections.

JEL classification: H63, C22
Keywords: fiscal policy sustainability, public debt, stationarity test, Markov Switching-ADF

1. Motivation and objectives

The COVID-19 pandemic, characterized by the rapid transmission of the virus across borders and lockdown restrictions, marked the largest global economic crisis in over a century [IMF 2022]. The pandemic created both supply and demand shocks, directly affecting government revenue due to unemployment, disrupted supply chains, and the bankruptcy of some institutions. Several countries provided stimulus packages to deal with the crisis, resulting in the largest one-year debt

* This article is a revised and shortened version of the undergraduate research paper of Premacio and Vidar. The original paper won the Gerardo P. Sicat Award for Best Undergraduate Research Paper for AY 2022-23 and is available at the UPSE Library. T. Monsod was the research adviser.

** Address all correspondence to dmpremacio@up.edu.ph and egvidar1@up.edu.ph.

surge since World War II. In addition, the war between Russia and Ukraine created inflationary pressure and uncertainty, disrupting global economic recovery [World Economic Outlook 2022]. This succession of events led to considerable swings in debt ratios [Gaspar et al. 2022], a highly uncertain fiscal policy environment [IMF 2022], and renewed interest in assessing the debt vulnerability and fiscal policy sustainability of countries as a basis for policy correction.

The empirical literature on fiscal policy sustainability largely involves examining whether the sovereign's present value borrowing constraint—whether the current value of public debt equals the discounted sum of future surpluses exclusive of interest payments—holds [Velinov 2015].[1] To do this, three methods have been applied: testing the stationarity of public debt and deficits, testing whether government revenues and expenditures, inclusive of interest payments, are suitably cointegrated, and testing whether a government's primary balance reacts positively to lagged increases in debt. Among these, the most common is the first, where many use unit root tests such as the Augmented Dickey-Fuller (ADF) test or Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test to determine the stationarity of the debt-to-GDP ratio. However, such standard tests are unreliable in the presence of nonlinearities and may lead to incorrect conclusions, prompting the use of regime-switching models, such as the Markov Switching ADF (MS-ADF).

Velinov [2015] points out that the conventional MS-ADF is, however, still restrictive—not all parameters are allowed to be state-dependent, no more than two states are considered, and higher-order autoregressions are left out. Velinov then introduces a "very general" MS-ADF, which "expands" existing models. This expanded MS-ADF allows an unrestricted number of lags, regime switches, and state-dependent parameters, thereby capturing irregularities in the depth and duration of phases. In an application to 16 OECD countries, this innovation is shown to be "an improvement on simpler existing models".

To our knowledge, the expanded MS-ADF has not yet been applied to countries in developing Asia. In fact, there are only a few studies that focus on Asia and most use samples that end before 2010, or prior to the global financial crisis. One exception is a working paper by Thuy [2018], which applied both stationarity and cointegration tests on the debt-to-GDP ratios of eight ASEAN countries from 1987 to 2017. Using time series difference-stationarity ADF tests, extensions of ADF for panel stationarity, and tests for panel cointegration, it was found that countries in the sample demonstrated fiscal sustainability, although in "weak" form.[2] A second exception is Bui [2019], which applied the fiscal reaction function method to a panel of 22 developing Asian countries, including the countries studied in Thuy [2018]. Bui [2019] found that, except for three countries, fiscal policy in the region was unsustainable.

[1] The rest of this paragraph draws heavily on Velinov [2015].

[2] As Thuy [2018] explains, weak sustainability means that the bubble term goes to zero at a slower rate, as long as the growth rate of debt does not exceed the growth rate of the economy.

The conflicting results from these two papers and our reading of Velinov [2015] motivate our research. How would results from the stationarity test approach compare with the fiscal reaction function approach if the expanded MS-ADF test of Velinov [2015], which improves on the methods used in Thuy [2018], is employed instead? In our view, this knowledge would be valuable in itself and can be a welcome addition to the limited literature on fiscal policy sustainability in developing Asia.

The rest of our paper is organized as follows. Section 2 provides more detail on the main methods used to examine fiscal sustainability, their shortcomings, and the expanded MS-ADF. Section 3 details our econometric model, estimation method and decision algorithm, while Section 4 discusses data, diagnostic tests and model selection per country. Section 5 presents the MS-ADF results and compares these with results from conventional stationarity tests and the fiscal reaction function model of Bui [2019]. Section 6 concludes.

2. Assessing fiscal policy sustainability and the expanded MS-ADF

Hamilton and Flavin [1986] argued that governments, like households, are subject to borrowing constraints. They discuss a government's present value borrowing constraint (PVBC), which means that the expected present value of expenditures, exclusive of interest payments, should not exceed the expected present value of receipts. A fiscal policy is deemed sustainable if it satisfies the PVBC.

Currently, three main methods are employed to determine whether the PVBC holds. These are summarized in Table 1, which draws heavily from discussions in Velinov [2015] and Bui [2019]. The first method is testing the stationarity of the first difference of the public debt stock, which is attributed to Hamilton and Flavin [1986]. Stationarity of public debt is a "sufficient condition" for fiscal policy sustainability, and, moreover, "as long as debt follows a stationary trajectory, it is sustainable regardless of its actual level" [Velinov 2015]. The second method can be viewed as equivalent to testing if debt is on a stationary trajectory and involves testing whether government receipts and government expenditures, inclusive of interest payments, are cointegrated with a vector $(1, -1)$ (Trehan and Walsh [1988;1991]; Hakkio and Rush [1991]).

The third method is from Bohn [1998;2007], who critiques the first two methods by showing why standard unit root and cointegration tests are "incapable of rejecting the consistency of data sets with the intertemporal budget constraint". Thus, "the common practice of judging a policy to be unsustainable on the basis of unit root and cointegration tests is invalid." Bohn suggests that examining the behavioral response of the primary balance may be a "more fruitful way of establishing debt sustainability". Specifically, a sufficient condition for the government to satisfy its intertemporal budget constraint is that the primary (non-interest) budget surplus is an increasing function of the (lagged) debt-GDP ratio.

TABLE 1. Main methods for assessing fiscal sustainability

Method	Developed by	Variables	How fiscal sustainability is determined	Criticisms
Stationarity Test	Hamilton and Flavin [1986]	debt-to-GDP ratio	Debt series is difference-stationary. Otherwise, it is unsustainable. Unit root tests, typically ADF or KPSS, are utilized.	The stationarity and cointegration conditions for sustainability are not necessary requirements for satisfying the intertemporal budget constraint. [Bohn 2007]. Conventional unit root tests have low power in the presence of nonlinearities and structural breaks which may lead to invalid conclusions (Afonso [2005]; Chen [2011]).
Cointegration Test	Trehan and Walsh [1988;1991] Hakkio and Rush [1991]	debt-to-GDP ratio government receipts government expenditures (including interest payments)	Government receipts and government expenditures, inclusive of interest payments, are cointegrated (with a coefficient of about 1). This is equivalent to testing if debt is on a stationary trajectory.	The same critique by Bohn [1998; 2007]. This assumes that expected real interest rate is constant, which is not always the case, and the null hypothesis of unit root is difficult to reject with short time series [Bui 2019].
Fiscal reaction function	Bohn [1998;2007]	primary balance and lagged public debt, controlling for temporary government expenditure, and the cyclical variations of output	The primary balance is an increasing function of the debt-to-GDP ratio.	This does not rule out a situation in which primary surpluses would need to exceed GDP to refinance debt [Ghosh et al. 2013]. In practice, lenders and policymakers are often concerned with perceived upper bounds of public debt which is not considered in this method [Velinov 2015].

The third method does not rule out the situation in which primary surpluses would need to exceed GDP to refinance public debt, however [Ghosh et al. 2013]. This is a problem in practice given that lenders and policymakers are often concerned with "perceived upper bounds on public debt", e.g., a limit on the debt-to-GDP ratio, say 60 percent, imposed by law or treaties [Velinov 2015]. Thus, notwithstanding the Bohn critique, many studies continue to use stationarity and cointegration test approaches.

The motivation to use regime-switching models to determine model stationarity is the fact that conventional stationarity tests such as ADF and KPSS do not always produce the same results. Velinov [2015] demonstrates this in an analysis of 16 OECD countries (Table 2). For instance, while both ADF and KPSS indicated that fiscal policy in Argentina and Finland was sustainable, they differed as regards Norway, Sweden, and UK (which KPSS deemed unsustainable) and Italy and Portugal (which ADF found to be unsustainable).

TABLE 2. Conflicting results of conventional stationarity tests

Unit Root Test	Sustainable	Unsustainable
Augmented Dickey-Fuller (ADF)	Argentina, Finland, Norway, Sweden, UK	France, Germany, Greece, Iceland, Ireland, Italy, Japan, Portugal, Spain, Switzerland, US
Kwiatkowski–Phillips–Schmidt–Shin (KPSS)	Argentina, Finland, Italy, Portugal	France, Germany, Greece, Iceland, Ireland, Japan, Norway, Spain, Sweden, Switzerland, UK, US

Source: Velinov [2015].

Velinov [2015] cites the low statistical power of conventional tests when the time series has a nonlinear nature [Chen 2011] as well as the bias of the ADF test towards nonrejection of the unit-root null hypothesis in the presence of structural breaks [Afonso 2005]. In contrast, the regime-switching MS-ADF (due to Hamilton [1989]) can accommodate nonlinearities, allow for states of nature (stationary and nonstationary) of public debt, and include the varying time paths of debt depending on states of nature associated with sources of systemic risks.

However, existing MS-ADF models still have shortcomings [Velinov 2015]. Most do not allow all parameters to be state-dependent, have not considered more than two states, and higher order autoregressions are often neglected, which may lead to erroneous conclusions. Thus, Velinov [2015] expanded existing models by allowing the number of lags and regimes to be unrestricted and parameters to be state-dependent.

Specifically, Velinov [2015] applied the following expanded MS-ADF model to test for unit roots:

$$\Delta B_t = \nu(S_t) + \Phi_1(S_t)B_{t\text{-}1} + \Phi_2(S_t)\Delta B_{t\text{-}1} + \Phi_3(S_t)\Delta B_{t\text{-}2} + \cdots + \Phi_{p+1}(S_t)\Delta B_{t\text{-}p+1} + u_t\,, \quad (1)$$

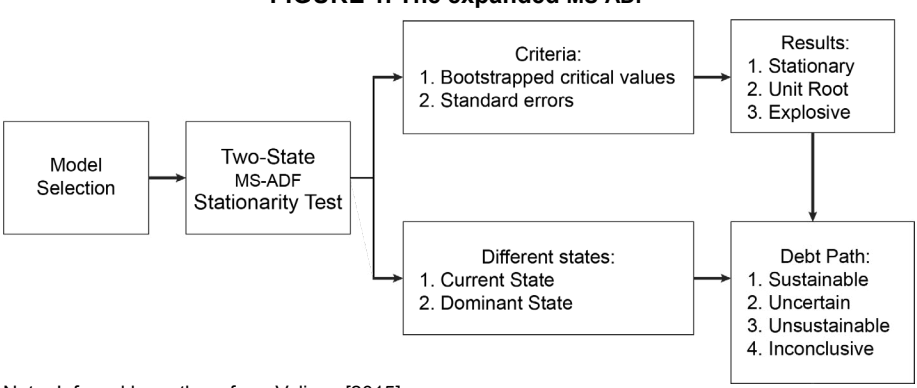
where B_t is government debt, Φ_1 is the coefficient on the first lag of B_t, and Φ_i, i = 2, … p +1, are coefficients on the first differences of government debt. S_t is a first order discrete-valued Markov process which can take on values 1..., M, allowing the numbers of lags and states to be unrestricted and all the parameters to be potentially state-dependent. The residual term u_t is assumed to have a normal distribution (as $u_t \sim$ Nid (0, $\boldsymbol{\sigma}^2$ (S_t)).

To select the appropriate number of lags of the model, Velinov [2015] used portmanteau tests based on MS-ADF residuals. The coefficient Φ_1 in (1) is assumed to be state-dependent, and all autoregressive coefficients of higher lag orders are allowed to switch. Further, the number of states, as well as whether a state is current, or whether it is dominant, is determined by examining the estimated smoothed probabilities of the countries. A state is considered current if it is the state of the last period of the sample. A state is dominant if it is the state which the country is in with the longest duration based on its fiscal policies. The current states are determined by observing the estimated smoothed probabilities of each country while the dominant states are determined by comparing state-transition probabilities.

The MS-ADF null hypothesis is that a unit root exists in each state ($\Phi_1(S_t)$ = 0, for S_t = 1...M). Negative values of test statistics imply stationarity while zero means that a unit root exists. Unlike conventional stationarity tests, positive values of the test statistic in each state can exist in the MS-ADF framework, indicating the presence of an explosive process.

After selecting the best model per country, some of which had two states and some three states, Velinov [2015] relied on both standard errors and parametrically bootstrapped critical values as criteria to determine the significance of parameters. In turn, the significance of parameters from both criteria determined whether the debt trajectory path of each country was "sustainable", "unsustainable", or "uncertain", i.e., it cannot be definitively categorized as sustainable or unsustainable. Whether a state was current and/or dominant, along with historical factors, also played a part in the assessment (Figure 1).

FIGURE 1. The expanded MS-ADF



Note: Inferred by authors from Velinov [2015].

Velinov [2015] did not provide a precise decision tree nor a list of rules to guide how exactly the two criteria could be used for judging the debt trajectory path of a country, however. For instance, it was not clear how conclusions are reached when only one criterion was satisfied (for instance, the parameter was significantly different from zero using standard errors only). Thus, to interpret our results, we construct a straightforward decision algorithm for using the two criteria, inferring rules from the analysis in Velinov [2015]. This is discussed in the next section.

3. Econometric model, estimation method, and decision rules

We adopt the model and approach in Velinov [2015].[3] Our econometric model is:

$$\Delta dgdp = v(S_t) + \Phi_1 (S_t) dgdp_{t-1} + \Phi_2 (S_t)\Delta dgdp_{t-1} + \Phi_3 (S_t)\Delta dgdp_{t-2} + \cdots + \Phi_{p+1}(S_t)\Delta dgdp_{t-p+1} + u_t \,, \quad (2)$$

where *dgdp* is government debt as a percentage of GDP. Scaling debt by GDP is necessary to avoid misleading results when performing unit root tests [Bohn 2019]. As in (1), Φ_1 is the coefficient on the first lag of B_t, and Φ_i, $i = 2, \ldots p + 1$, are coefficients on the first differences of government debt; S_t is first order discrete-valued Markov process which can take on values 1..., M, and the residual term u_t is assumed to be normally distributed.

We first undertake conventional ADF and KPSS unit root tests to determine model stationarity. The Akaike's Information Criteria (AIC) and Schwarz Information Criterion (SBIC) are used to select the best lag lengths; a lower AIC and SBIC indicate a better fit. In cases where the optimal lag order is different for AIC and SBIC, both lags were utilized. Finding the appropriate number of lags is important because selecting higher lag orders can increase the mean-square forecast errors of the VAR while underfitting the lag length often generates autocorrelated errors [Lütkepohl 1993].

Inconsistent ADF and KPSS results motivate the regime-switching MS-ADF. We check for serial autocorrelation (using residual portmanteau and LM tests), homogeneity (using ARCH-LM) and structural breaks (using Chow, Recursive and CUSUM-SQ tests) to support the use of a regime-switching model. The portmanteau test based on MS-ADF residuals is also used to choose the optimal lag among those with multiple lag orders.

We rely on two criteria to determine parameter significance—standard errors and bootstrapped critical values—to assess the sustainability of debt trajectories. To bootstrap, a non-parametric bootstrapping program is used, which offers a robust alternative and makes fewer assumptions compared to classic parametric

[3] Unless otherwise indicated, the methods described in this section follow Velinov [2015], which was explained in the preceding section.

methods.[4] We then apply a straightforward decision algorithm for using the two criteria, which we construct based on our interpretation of the Velinov's analysis. This algorithm is designed for a two-state model (since all countries in our sample were determined to have a two-state model) and has six rules. The rules are:

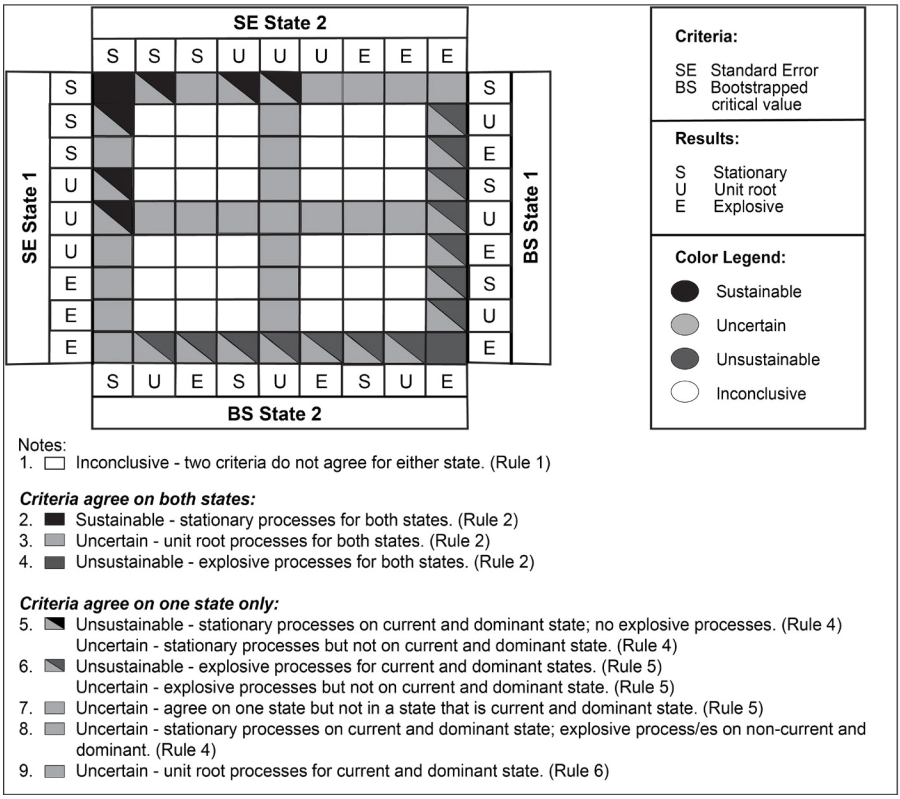
> Rule 1: The results from both criteria must be the same for either state 1 or state 2. Otherwise, the results are inconclusive, i.e., no conclusion can be drawn as to the sustainability of the debt path.
>
> Rule 2: If the two criteria agree in both states, then the debt path is sustainable if the criteria indicate stationary processes in both states; unsustainable if the criteria indicate explosive processes in both states; and uncertain if the criteria indicate unit root processes in both states, or the criteria indicate stationary processes in one state and explosive processes in the other state.
>
> Rule 3: When the two criteria agree for either state 1 or state 2 only, they must agree in a state that is both current and dominant. Otherwise, the result is uncertain.
>
> Rule 4: If Rule 3 holds, the debt path is sustainable if both criteria indicate a stationary process for that state which is current and dominant, and neither criterion indicates an explosive process for the other non-current and non-dominant state. If the latter does not hold, and a criterion indicates an explosive process for the other state, then the debt path is uncertain.
>
> Rule 5: If Rule 3 holds, the debt path is unsustainable if both criteria indicate an explosive process for the state that is current and dominant.
>
> Rule 6: If Rule 3 holds, the debt path is uncertain if both criteria indicate unit root processes for the state that is current and dominant.

The matrix in Figure 2 summarizes the algorithm. To provide an example, say state 1 is stationary in both criteria and state 2 is uncertain for both criteria (S-U-S-U).[5] State 1 is both current and dominant. Using the matrix, and moving from the left (S), to the top (U), to the right (S), and bottom (U), we see that S-U-S-U intersects at a block which indicates either "sustainable" or "uncertain" (row 1, column 5). Since Rule (2) applies—i.e., state 1 is stationary in both criteria and is also the current state and the dominant state—the debt path is sustainable.

[4] This replaces the parametric method of Psaradakis [1998] which was employed by Velinov [2015].

[5] (S-U-S-U) follows the format (State 1 for criterion 1 - State 2 for criterion 1 - State 1 for criterion 2 - State 2 for criterion 2).

FIGURE 2. MS-ADF fiscal sustainability decision matrix



The statistical software used in this study is Stata 17 MP-Parallel Edition and R version 4.1.3 (March 10, 2022). Most of the diagnostic tests and the MS-ADF test are conducted in Stata. The portmanteau test was done using R.

4. Data, diagnostic tests, and model selection per country

4.1. Data

We use annual frequency data on the (nominal) gross debt-to-GDP ratio taken from the World Economic Outlook Database, covering 1999–2017, for 22 developing Asian economies. We use a data set identical to Bui [2019] to produce comparable results.

As observed from Figure 3, most of the countries kept their debt below 60 percent of their GDP during the period covered. Among the 22 countries, Myanmar experienced the highest debt-to-GDP ratio (at 252 percent) in 1999 to 2007, Bhutan in 2011 to 2017, India in 2008, and Sri Lanka, from 2009 to 2010. On the other hand, Kiribati remained to have the lowest debt-to-GDP ratio in the region until 2014.

One can also broadly identify the stationarity of a series using visualization. Debt-to-GDP is stationary if it does not significantly change over time; graphically, a stationary trajectory would look like a straight line. For instance, Bangladesh, Cambodia, India, Kiribati, Marshall Islands and Micronesia show no visible long-term trend and thus, it is possible that these countries have stationary debt processes and sustainable debt paths (Figure 4). On the other hand, Indonesia, Lao PDR, Myanmar, Nepal, Papua New Guinea, Philippines, Sri Lanka, Solomon Islands, Thailand, and Vietnam appear to have trajectories on a downward trend (Figure 5) while Bhutan, China, Malaysia, Maldives, and Vanuatu exhibit trajectories on an upward trend (Figure 6). It is possible that these countries have nonstationary processes.

FIGURE 3. Time plot of debt-to-GDP ratios per country

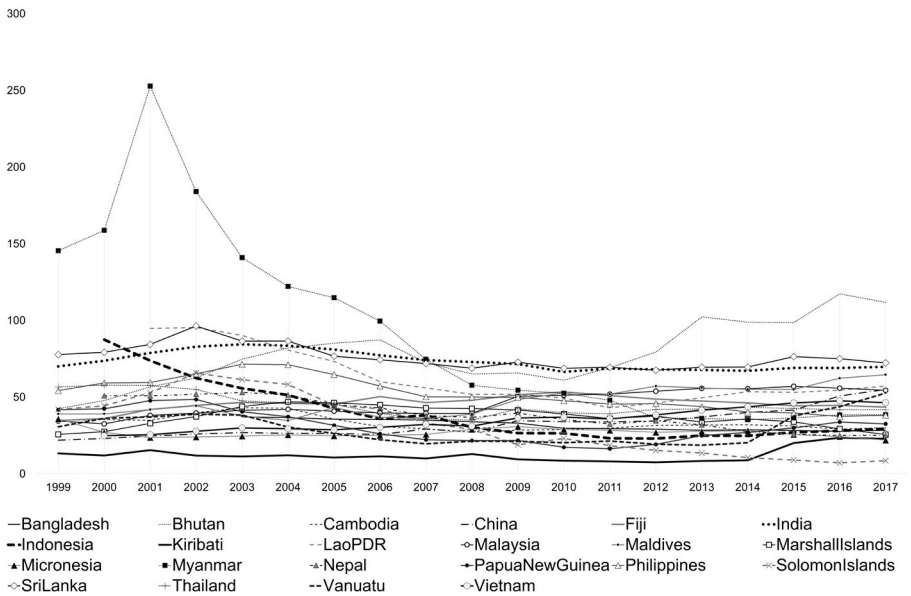


FIGURE 4. Debt-to-GDP Graph of Bangladesh, Cambodia, India, Kiribati, Marshall Islands, and Micronesia

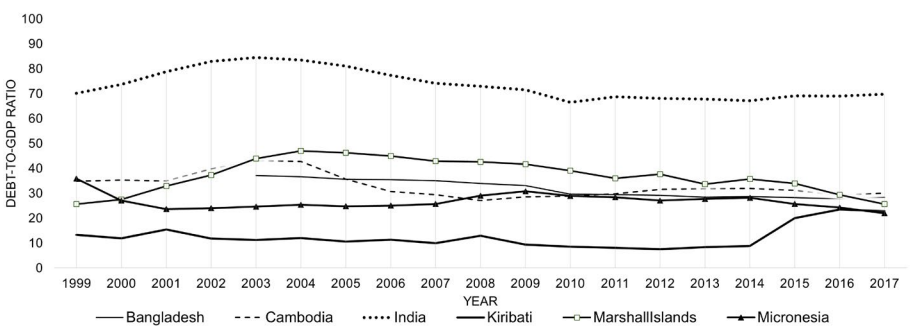


FIGURE 5. Debt-to-GDP Graph of Indonesia, Lao PDR, Myanmar, Nepal, Papua New Guinea, Philippines, Sri Lanka, Solomon Islands, and Thailand

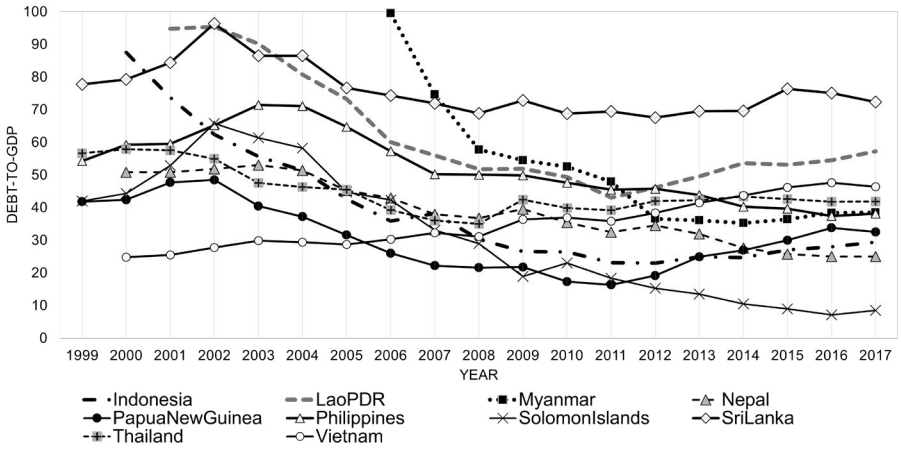
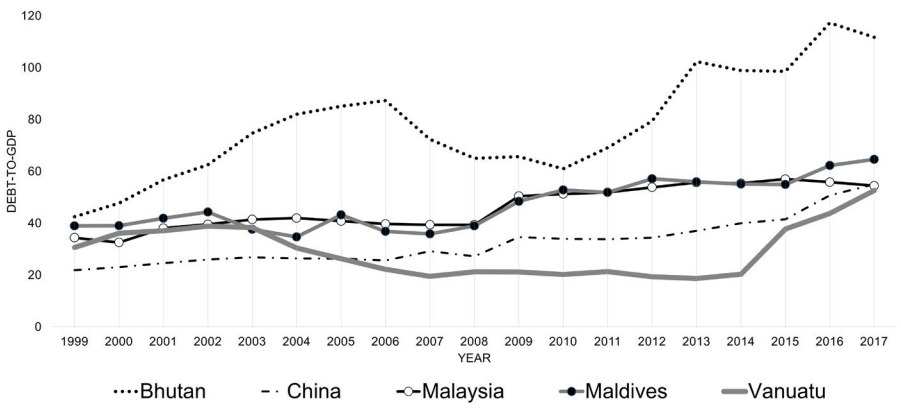


FIGURE 6. Debt-to-GDP Graph of Bhutan, China, Malaysia, Maldives, and Vanuatu



4.2. Diagnostic tests

Table 3 presents the results of the diagnostic tests conducted. For conventional unit root tests, lag length is determined from the AIC and SBIC tests. If these tests provide different results, multiple lag lengths are used for the ADF and KPSS.

The debt trajectories of 11 countries are found to be stationary using only one of the tests but not the other. Specifically, Fiji, Indonesia, Lao PDR, Myanmar, and Papua New Guinea are considered stationary using the ADF test only, while Cambodia, Malaysia, Micronesia, Nepal, Philippines, and Solomon Islands are considered stationary by KPSS only. Only India and Sri Lanka's debt trajectories were found to be stationary by both the ADF and KPSS tests. The inconsistency of results from the two tests motivates the use of the MS-ADF.

TABLE 3. Diagnostic tests for all countries

Country	Lag Length	Stationarity Tests		Autocorrelation Tests+				Heteroske-dasticity Test+
		ADF*	KPSS**	Q_{12}[1]	Q^{A}_{12}[2]	LM_5[3]	LMF_5[4]	$ARCH_{LM}$ (12)[5]
Bangladesh	1	-1.38	0.15	0.99	0.89	0.65	0.67	0.90
Bhutan	1	-1.31	0.12	0.87	0.51	0.45	0.49	0.34
Cambodia	2	-1.60	0.08	0.15	0.02	0.65	0.66	0.26
China	2	2.12	0.17	0.46	0.04	0.20	0.29	0.89
Fiji	1	-3.13	0.23	0.99	0.89	1.00	1.00	0.55
India	5	-3.03	0.10	0.16	0.00	0.13	0.35	0.56
	1	-1.56	0.12	0.29	0.04	0.12	0.21	0.98
Indonesia	4	-0.98	0.15	0.35	0.01	0.26	0.41	0.32
	1	-3.62	0.25	0.82	0.49	0.42	0.47	0.35
Kiribati	1	-0.92	0.18	0.99	0.92	0.26	0.33	0.77
Lao PDR	1	-3.26	0.23	0.87	0.35	0.69	0.69	0.22
Malaysia	1	-1.48	0.08	1.00	0.95	0.87	0.86	0.98
Maldives	1	-0.23	0.17	0.84	0.41	0.71	0.71	0.37
Marshall Islands	1	-1.50	0.23	0.60	0.18	0.10	0.18	0.55
Micronesia	2	-1.15	0.10	0.02	0.00	0.29	0.37	0.85
Myanmar	4	-3.82	0.12	0.99	0.89	0.07	0.23	0.29
	3	-2.50	0.13	0.99	0.91	0.47	0.52	0.92
Nepal	5	-1.13	0.14	0.71	0.09	0.06	0.35	0.30
	1	-0.53	0.06	0.97	0.75	0.31	0.38	0.16
Papua New Guinea	4	-2.37	0.13	0.02	0.00	0.77	0.76	0.75
	2	-2.76	0.17	0.06	0.00	0.05	0.14	0.57
Philippines	2	-0.89	0.10	0.28	0.07	0.22	0.31	0.70
Solomon Islands	3	-2.54	0.09	0.60	0.26	0.03	0.14	0.95
Sri Lanka	3	-4.12	0.10	0.67	0.19	0.01	0.10	0.17
Thailand	1	-2.29	0.23	0.81	0.39	0.94	0.93	0.32
Vanuatu	3	-1.57	0.14	0.22	0.01	0.27	0.37	0.42
Vietnam	2	-0.18	0.15	0.74	0.25	0.39	0.45	0.79

*Critical values are -3.75 at one percent, -3 at five percent, -2.63 at ten percent.
**Critical values are 0.216 at one percent, 0.176 at 2.5 percent, 0.146 at five percent and 0.119 at ten percent.
+ Only *p*-values are reported.
[1] Portmanteau test statistic using 12 lags with a χ^2 distribution.
[2] Adjusted portmanteau test statistic using 12 lags with a χ^2 distribution.
[3] LM test statistic using five lags with a χ^2 distribution.
[4] LM test statistic using five lags with an *F* distribution.
[5] ARCH-LM test statistic using 12 lags with a χ^2 distribution.

The next four columns contain the p-values of portmanteau and LM autocorrelation tests; columns 6 and 8 are the results when the tests were adjusted to accommodate smaller sample sizes. Most countries cannot reject the null hypothesis of no residual autocorrelation. The last column shows the p-values of the ARCH LM test for heteroskedasticity. For all countries, the null of no conditional heteroskedasticity cannot be rejected.

Chow, recursive and CUSUM-SQ tests (not reported in the table but available upon request) indicate evidence of structural breaks in the time series of majority of the countries, further supporting the use of a regime-switching model.

4.3. Model selection per country

For each country, we select the appropriate number of lags of the regime-switching model using the portmanteau test, while the number of states is determined by observing the estimated smoothed probabilities (shown in the Appendix and explained further in Section 5). Notably all countries are found to have two states.

Table 4 presents the MS-ADF model used for each country, where MS(M) stands for Markov switching with M states, ADF(p) for ADF model with p lags, A for switching autoregressive parameters, and H for a switching variance parameter. As earlier mentioned, the coefficient of the first lag of government debt, Φ_1, is assumed to be state-dependent while the autoregressive coefficients of higher lag orders are allowed to switch. Parameter stability tests were conducted to check if variance is state-dependent.

TABLE 4. Model selected per country

Country	Model	Country	Model	Country	Model
Bangladesh	MS (2)-ADF (1) AH	Lao PDR	MS (2)-ADF (1)A	Papua New Guinea	MS (2)-ADF (2)A
Bhutan	MS (2)-ADF (1) AH	Malaysia	MS (2)-ADF (1)A	Philippines	MS (2)-ADF (1)A
Cambodia	MS (2)-ADF (2)A	Maldives	MS (2)-ADF (1) AH	Solomon Islands	MS (2)-ADF (3)A
China	MS (2)-ADF (2)A	Marshall Islands	MS (2)-ADF (1)A	Sri Lanka	MS (2)-ADF (3)A
Fiji	MS (2)-ADF (1) AH	Micronesia	MS (2)-ADF (2)A	Thailand	MS (2)-ADF (1)A
India	MS (2)-ADF (5)A	Myanmar	MS (2)-ADF (3)A	Vanuatu	MS (2)-ADF (3) AH
Indonesia	MS (2)-ADF (4)A	Nepal	MS (2)-ADF (5) AH	Vietnam	MS (2)-ADF (1) AH
Kiribati	MS (2)-ADF (1)A				

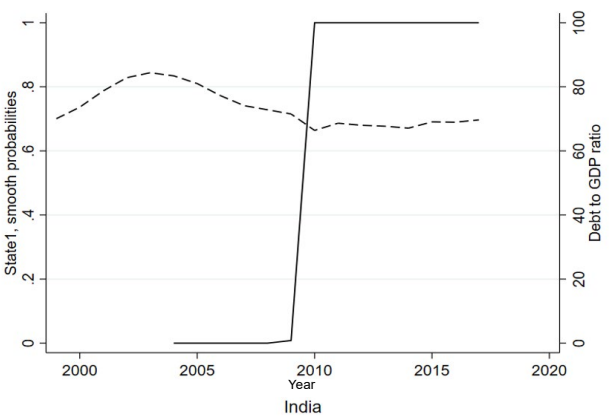
5. MS-ADF empirical results

Table 5 shows estimated parameters of state 1 and state 2 with their standard errors and bootstrapped critical values. Significant negative coefficients of state 1 and state 2 of the MS-ADF model—that is, ($\Phi(m) < 0$) which are significant at the ten percent level based on standard errors and bootstrapped critical values—indicate stationary states, while significant positive coefficients—i.e., ($\Phi(m) > 0$) which are significant at the ten percent level—indicate explosive states. The countries are arranged based on the coefficient Φ_1—smallest to largest. In other words, from the most stationary to least stationary based on their first state.

Further, Table 5 presents parameters $p(nm)$ which are the probabilities of switching from state n to state m. These parameters are used to determine which is the dominant state. Specifically, if $p_{11} > p_{22}$, then state 1 is the dominant state. The dominant state along with the current state are essential in interpreting the results of the MS-ADF (as evident in our decision algorithm discussed in Section 3).

To illustrate, Figure 7 shows the smoothed probabilities of state 1 of India.[6] The solid line is the smoothed probabilities (left axis) while the dashed line is the debt-to-GDP ratio (right axis). The initial state is state 1 but if the probability rises to 1, it implies that it transitioned to state 2; if the probability drops back to 0, it returns to state 1. For India, the initial state is state 1. It then transitioned to state 2 in 2009 and remained there. Hence, India's current state is state 2 and its dominant state is state 1 (0.88 > 0.29).

FIGURE 7. State 1, smoothed probabilities, India



[6] In a two-state model, the smoothed probabilities of state 2 are a mirror image of state 1's. Thus, graphs for state 2 need not be shown.

TABLE 5. State parameter estimates and state transition probabilities

Countries	$\Phi 1$	SE	BS	$\Phi 2$	SE	BS	p_{11}	SE	p_{22}	SE
Maldives	-0.39¶ˣ»	0.11	0.45	1.13*	0.10	0.16	0.72	0.20	0.06	0.07
Micronesia	-0.12ˣ»	0.04	0.35	0.30¶	0.41	0.51	0.92	0.09	0.18	0.11
Sri Lanka	-0.14ˣ»	0.69	0.42	-0.38¶ ˣ	0.07	1.85	0.90	0.26	0.37	0.41
Cambodia	0.17ˣ»	0.23	0.27	-0.68¶ˣ	0.21	0.97	0.94	0.04	0.21	0.15
India	0.03»	0.14	0.29	-0.79¶ ˣ	0.05	0.51	0.88	0.08	0.29	0.25
Papua New Guinea	0.50¶»	0.35	0.35	0.65**	0.14	0.26	0.83	0.07	0.17	0.09
Indonesia	0.38*»	0.19	0.38	0.35¶**	0.03	0.21	0.68	0.24	0.39	0.56
Kiribati	0.57*»	0.17	0.46	0.13¶	0.70	14.57	0.95	0.06	0.12	0.09
Vanuatu	1.04**»	0.26	0.45	-0.46ˣ	0.23	1.62	0.86	0.09	0.16	0.23
Nepal	0.40¶*»	0.05	0.42	0.44*	0.07	0.31	0.88	0.07	0.11	0.08
Solomon Islands	0.74**»	0.08	0.53	1.20¶	0.09	0.74	0.70	0.42	0.07	0.05
Malaysia	0.44¶**»	0.04	0.32	0.90*	0.31	0.33	0.86	0.10	0.08	0.06
Lao PDR	0.95¶**»	0.08	0.43	0.67**	0.08	0.26	0.84	0.34	0.16	0.14
Bhutan	0.81¶**	0.05	0.40	1.02**»	0.10	0.15	0.51	0.21	0.72	0.30
China	1.15¶**»	0.05	0.36	1.58**	0.06	0.77	0.73	0.10	0.18	0.15
Thailand	1.01**»	0.04	0.37	0.55¶**	0.04	0.24	0.74	0.21	0.13	0.07
Myanmar	0.41¶**»	0.01	0.35	1.24*	0.01	0.31	0.88	0.07	0.54	0.56
Philippines	1.14**»	0.02	0.32	0.78¶**	0.04	0.57	0.92	0.18	0.05	0.04
Fiji	0.81¶**	0.01	0.32	0.59*»	0.13	0.41	0.00	0.00	0.22	0.11
Marshall Islands	1.00¶**»	0.00	0.30	1.00**	0.00	0.38	0.94	0.05	0.15	0.40

Note: $\Phi(m)$ the state parameter, $p(nm)$ the transition probability of state n to m, SE refers to the standard error of the parameter $\Phi(m)$, BS is the bootstrapped critical value of $\Phi(m)$.
No parameter estimates are obtained for Bangladesh and Vietnam hence their exclusion from this table.
¶ The current state
» The dominant state
ˣ Stationary according to one criterion
ˣˣ Stationary according to both criteria
* Explosive according to one criterion
** Explosive according to both criteria

Table 6 summarizes the results and debt trajectory of each country. It shows that Bhutan, China, Lao PDR, Marshall Islands, Philippines, and Thailand have explosive processes for both states using both standard errors and bootstrapped critical values. By Rule 2 of the decision algorithm, this implies that their debt path is unsustainable.

For Vanuatu, state 1 is characterized as explosive by both criteria while state 2 is classified as stationary and unit root. Vanuatu's state 1 is current and dominant, hence by Rule 5, its debt trajectory is unsustainable.

Cambodia, Fiji, India, Indonesia, Kiribati, Malaysia, Maldives, Micronesia, Myanmar, Papua New Guinea, and Sri Lanka have uncertain debt paths. This is according to either Rule 2, where criteria indicate unit root processes for both states; Rule 3, where the current state and dominant state is not the same, or Rule 6, where criteria agree in only one state that is current and dominant and which has a unit root process. These scenarios make it difficult to evaluate whether their fiscal policies satisfy the PVBC condition, and to definitively categorize their debt path as sustainable or unsustainable. For example, Papua New Guinea has state 1 with unit root process and state 2 with an explosive process which implies that their debt path can either be unsustainable or uncertain. Since the state with a unit root process is current and dominant, Papua New Guinea's debt path is uncertain.

Two countries—Nepal and Solomon Islands—have inconclusive results by Rule 1. This means that the standard errors and bootstrapped critical values did not agree with their assessment of state 1 and state 2. Another two countries—Bangladesh, and Vietnam—are considered inconclusive because parameters could not be estimated. This could be because they have missing debt-to-GDP data leading to smaller sample sizes compared to other countries.

Table 7 presents the MS-ADF results alongside results from the fiscal reaction function of Bui [2019]. Out of 22 countries in the sample, the MS-ADF and Bui [2019] agree that six (Bhutan, China, Lao, Marshall Islands, Thailand, and Vanuatu) have debt trajectories that are unsustainable but disagree on all the rest. In particular, the debt paths of nine others (Cambodia, Fiji, Kiribati, Malaysia, Maldives, Micronesia, Myanmar, Papua New Guinea, and Sri Lanka) deemed by Bui [2019] to be unsustainable are considered to be uncertain using MS-ADF. Many of these countries with uncertain debt paths as per MS-ADF have at least one explosive state, which may imply that at some point, their fiscal policy may be unsustainable. The fiscal reaction function method of Bui [2019], however, does not allow for a regime switch and may have automatically categorized these countries as having unsustainable debt paths although it is not yet definite. Bangladesh, Nepal, Solomon Islands, and Vietnam, which were deemed unsustainable in Bui [2019], had inconclusive results using MS-ADF.

TABLE 6. Debt trajectory per country

Countries*	State 1 SE	State 2 SE	State 1 BS	State 2 BS	Current State	Dominant State	Debt Trajectory Path
Bhutan	Explosive	Explosive	Explosive	Explosive	State 1	State 2	Unsustainable
Cambodia	Unit Root	Stationary	Unit Root	Unit Root	State 2	State 1	Uncertain
China	Explosive	Explosive	Explosive	Explosive	State 1	State 1	Unsustainable
Fiji	Explosive	Explosive	Explosive	Unit Root	State 1	State 2	Uncertain
India	Unit Root	Stationary	Unit Root	Unit Root	State 2	State 1	Uncertain
Indonesia	Explosive	Explosive	Unit Root	Explosive	State 2	State 1	Uncertain
Kiribati	Explosive	Unit Root	Unit Root	Unit Root	State 2	State 1	Uncertain
Lao PDR	Explosive	Explosive	Explosive	Explosive	State 1	State 1	Unsustainable
Malaysia	Explosive	Explosive	Unit Root	Explosive	State 1	State 1	Uncertain
Maldives	Stationary	Explosive	Unit Root	Explosive	State 1	State 1	Uncertain
Marshall Islands	Explosive	Explosive	Explosive	Explosive	State 1	State 1	Unsustainable
Micronesia	Stationary	Unit Root	Unit Root	Unit Root	State 2	State 1	Uncertain
Myanmar	Explosive	Explosive	Unit Root	Explosive	State 1	State 1	Uncertain
Nepal	Explosive	Explosive	Unit Root	Unit Root	State 1	State 1	Inconclusive
Papua New Guinea	Unit Root	Explosive	Unit Root	Explosive	State 1	State 1	Uncertain
Philippines	Explosive	Explosive	Explosive	Explosive	State 2	State 1	Unsustainable
Solomon Islands	Explosive	Explosive	Unit Root	Explosive	State 1	State 1	Inconclusive
Sri Lanka	Unit Root	Stationary	Unit Root	Unit Root	State 2	State 1	Uncertain
Thailand	Explosive	Explosive	Explosive	Explosive	State 2	State 1	Unsustainable
Vanuatu	Explosive	Stationary	Explosive	Unit Root	State 1	State 1	Unsustainable

* Note: Bangladesh and Vietnam are not listed here, as they were not in Table 5.

TABLE 7. Comparative results: expanded MS-ADF and Bui [2019]

Countries	*ADF*	*KPSS*	Expanded MS-ADF	Fiscal reaction function of Bui [2019]
Bangladesh	*Unsustainable*	*Unsustainable*	Inconclusive	Unsustainable
Bhutan	*Unsustainable*	*Unsustainable*	Unsustainable	Unsustainable
Cambodia	*Unsustainable*	*Sustainable*	Uncertain	Unsustainable
China	*Unsustainable*	*Unsustainable*	Unsustainable	Unsustainable
Fiji	*Sustainable*	*Unsustainable*	Uncertain	Unsustainable
India	*Sustainable*	*Sustainable*	Uncertain	Sustainable
Indonesia	*Sustainable*	*Unsustainable*	Uncertain	Sustainable
Kiribati	*Unsustainable*	*Unsustainable*	Uncertain	Unsustainable
Lao PDR	*Sustainable*	*Unsustainable*	Unsustainable	Unsustainable
Malaysia	*Unsustainable*	*Sustainable*	Uncertain	Unsustainable
Maldives	*Unsustainable*	*Unsustainable*	Uncertain	Unsustainable
Marshall Islands	*Unsustainable*	*Unsustainable*	Unsustainable	Unsustainable
Micronesia	*Unsustainable*	*Sustainable*	Uncertain	Unsustainable
Myanmar	*Sustainable*	*Unsustainable*	Uncertain	Unsustainable
Nepal	*Unsustainable*	*Sustainable*	Inconclusive	Unsustainable
Papua New Guinea	*Sustainable*	*Unsustainable*	Uncertain	Unsustainable
Philippines	*Unsustainable*	*Sustainable*	Unsustainable	Sustainable
Solomon Islands	*Unsustainable*	*Sustainable*	Inconclusive	Unsustainable
Sri Lanka	*Sustainable*	*Sustainable*	Uncertain	Unsustainable
Thailand	*Unsustainable*	*Unsustainable*	Unsustainable	Unsustainable
Vanuatu	*Unsustainable*	*Unsustainable*	Unsustainable	Unsustainable
Vietnam	*Unsustainable*	*Unsustainable*	Inconclusive	Unsustainable

Source: Authors' computations except for last column which are from models 1 and 2 of Bui [2019].

Of the three countries Bui [2019] found to have sustainable debt paths, two were found by MS-ADF to have unsustainable debt paths (Indonesia, Philippines) and one was found to be uncertain (India). This might be because while the three countries have significantly reduced their debt-to-GDP ratios over the years, whether adjustments are sufficient to allow them to easily refinance their debt is not clear. The model in Bui [2019] would have captured the former but not the latter since the fiscal reaction function only considers how fiscal surpluses react to changes in debt and does not consider the actual fiscal position. The smoothed probabilities of the MS-ADF model also show that among the three countries, only India had a regime switch after the 2008 global recession, possibly explaining the "uncertain" (rather than "unsustainable") finding. In contrast, Bui [2019] found that all three countries adjusted their fiscal policies after 2008.

Table 7 also presents the results from the conventional ADF and KPSS tests (in columns 2 and 3). We note that the MS-ADF provides clarity, nuance, and likely more accuracy (as expected) relative to results from ADF and KPSS. Of the 11 countries mentioned in section 4.2 as having conflicting ADF and KPSS results, the MS-ADF classifies seven as having uncertain debt paths (Cambodia, Fiji, Indonesia, Malaysia, Micronesia, Myanmar and Papua New Guinea), two as having unsustainable debt paths (Lao, Philippines), and two as having inconclusive results (Nepal, Solomon Islands). For the 11 others that did not have different ADF and KPSS results, the MS-ADF concurs in only five instances (Bhutan, China, Marshall Islands, Thailand and Vanuatu, all classified unsustainable).[7] The other six were found by MS-ADF to have either inconclusive or uncertain results, although these countries were found to have either unsustainable (Bangladesh, Kiribati, Maldives, Vietnam) or sustainable (India, Sri-Lanka) debt trajectories by both ADF and KPSS.

6. Concluding remarks

This paper sets out to measure fiscal policy sustainability in 22 countries in developing Asia using a never-before applied regime-switching stationarity test, the expanded MS-ADF, due to Velinov [2015]. The model has better statistical power in the presence of nonlinearities and structural breaks, addressing the weaknesses of conventional stationarity tests like the ADF and KPSS. We use a data set that will allow a comparison of results with Bui [2019], which uses a fiscal reaction function approach to assess fiscal policy sustainability.

In contrast to results in Bui [2019], who finds that all countries have either unsustainable (19 countries) or sustainable (three countries) fiscal policies, the MS-ADF indicates a more nuanced picture of the region. Eleven countries are classified as having uncertain fiscal sustainability (Cambodia, Fiji, India, Indonesia, Kiribati, Malaysia, Maldives, Micronesia, Myanmar, Papua New Guinea, and Sri Lanka), while just seven are classified as having unsustainable fiscal policies (Bhutan, China, Lao PDR, Marshall Islands, Philippines, Thailand, and Vanuatu). Results for the remaining four countries (Bangladesh, Nepal, Solomon Islands and Vietnam) are inconclusive; no country is found to have sustainable fiscal policies.

In other words, the MS-ADF concurs with Bui [2019] for only 6 of the countries the latter found to be unsustainable, reclassifying 13 others as either uncertain or inconclusive. It also reclassifies one country that Bui [2019] found to be sustainable as uncertain. We also note that, as expected, the MS-ADF provides more clarity and nuance to the results arising from conventional stationarity tests.

[7] Interestingly, our ADF results are different from Thuy [2018]. For instance, our ADF did not find the debt path of Thailand to be sustainable, nor Indonesia and Myanmar's to be unsustainable. We suspect this is because Thuy [2018] failed to determine the appropriate lag length in conducting the ADF method.

Using the MS-ADF, it is possible that debt trajectories are not definitively sustainable or unsustainable but somewhere in-between. This nuance may be helpful in identifying areas of fiscal policy to improve and prioritize, leading to more suitable policy corrections.

The MS-ADF model encounters problems in finding the optimal solution for data with missing values and small sample size. For future studies, we recommend the use of a longer time series per country. The implications of the Bohn [1998, 2007] critique, about the validity of the stationarity conditions, also needs to be thought through vis the MS-ADF framework, notwithstanding the claims by Velinov [2015] that the MS-ADF is a practical choice.

References

Afonso, A. [2005] "Fiscal sustainability: the unpleasant European case", *Finanz Archiv/Public Finance Analysis* 61(19):19-44.

Bohn, H. [1998] "The behavior of US public debt and deficits", *Quarterly Journal of Economics* 113(3):949-963.

Bohn, H. [2007] "Are stationarity and cointegration restrictions really necessary for the intertemporal budget constraint?", *Journal of Monetary Economics* 54(7):1837-1847.

Bui, D.T. [2019] "Fiscal sustainability in developing Asia: new evidence from panel correlated common effect model", *Journal of Asian Business and Economic Studies* 27(1):66–80. https://doi.org/10.1108/jabes-01-2019-0001.

Chen, S. [2011] "Current account deficits and sustainability: evidence from the OECD countries", *Economic Modelling* 28(4):1455-1464.

Ghosh A., J. Kim, E. Mendoza, J.D. Ostry, and M. Qureshi [2013] "Fiscal fatigue, fiscal space and debt sustainability in advanced economies", *Economic Journal* 123(566):F4-F30.

Hakkio, C.S. and M. Rush [1991] "Is the budget deficit 'too large?'", *Economic Inquiry* 29(3):429-445.

Hamilton, J.D. and M.A. Flavin [1986] "On the limitations of government borrowing: a framework for empirical testing", *The American Economic Review* 76(4):808–819. http://www.jstor.org/stable/1806077. Accessed March 16, 2023.

Haug, A.A. [1991] "Cointegration and government borrowing constraints: evidence for the United States", *Journal of Business & Economic Statistics* 9(1):97-101.

International Bank for Reconstruction and Development/The World Bank [2022] "The economic impacts of the COVID-19 crisis", in *World development report 2022: finance for an equitable recovery*. Washington, DC: The World Bank.

International Monetary Fund (IMF) [2022] *Fiscal monitor: fiscal policy from pandemic to war*. Washington, DC: IMF.

Lütkepohl, H. [1993] *Introduction to multiple time series analysis* (2nd ed.). Berlin: Springer. https://doi.org/10.1007/978-3-642-61695-2.

Psaradakis, Z. [1998] "Bootstrap-based evaluation of Markov switching time series models", *Econometric Reviews* 17(3):275-288.

Smith, A.D., P.A Naik, and C. Tsai [2005] "Markov-switching model selection using Kullback-Leibler divergence", *UC Davis Agricultural and Resource Economics Working Paper No*. 05-001

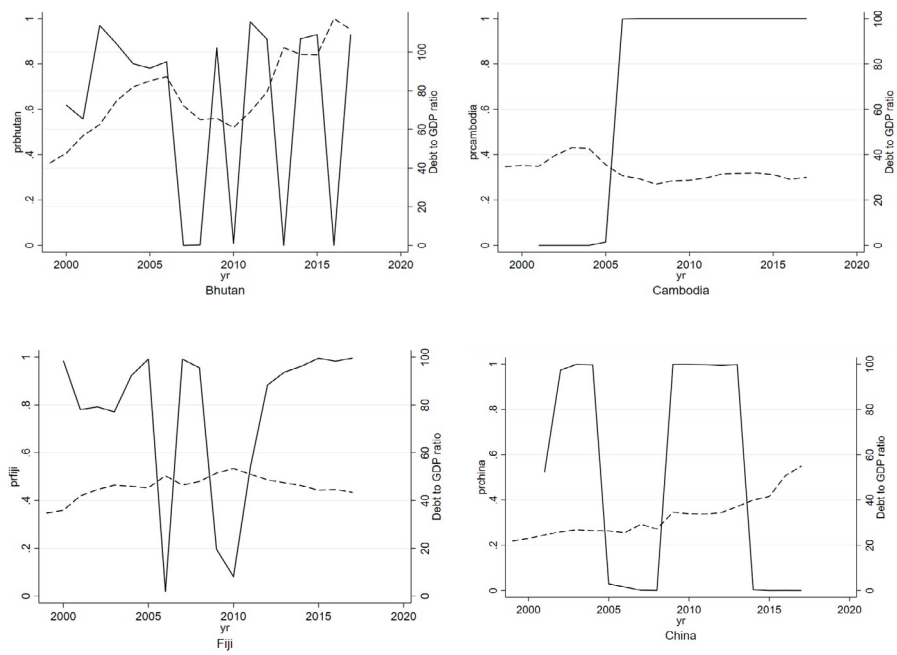
Tanner, E. and P. Liu [1994] "Is the budget deficit 'too large'? Some further evidence", *Economic Inquiry* 32(3):511-518.

Thuy, N.T. [2018] "Fiscal sustainability in some ASEAN countries: evidence from a panel data analysis", *Japanese Ministry of Finance*. https://www.mof.go.jp/pri/international_exchange/visiting_scholar_program/Nguyen2.pdf.

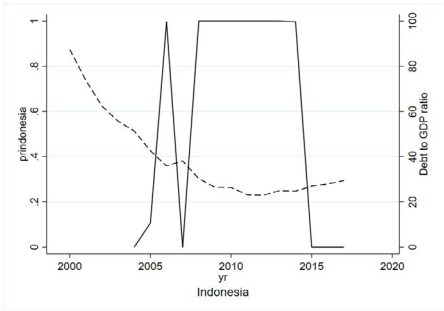
Trehan, B. and C.E. Walsh [1991] "Testing intertemporal budget constraints: theory and applications to US federal budget and current account deficits", *Journal of Money, Credit, and Banking* 23(2):206–223. https://doi.org/10.2307/1992777. Accessed 16 Mar. 2023.

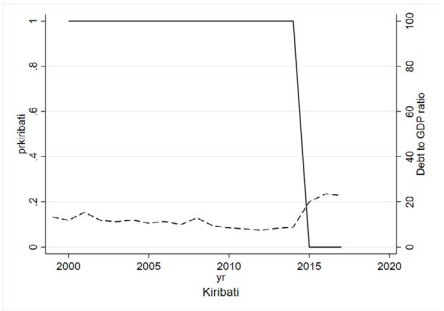
Velinov, A. [2015] "Assessing fiscal-policy sustainability: on the different states of the debt-to-GDP process", *Finanz Archiv/Public Finance Analysis* 71(4): 415–39. http://www.jstor.org/stable/24807743. Accessed March 6, 2023.

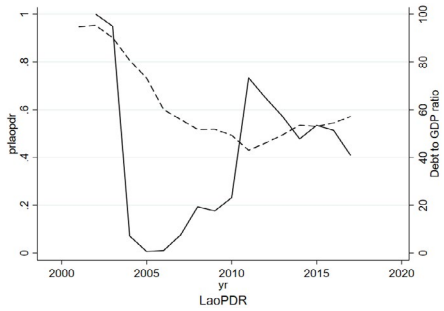
Appendix. Model estimated smoothed probabilities of State 1

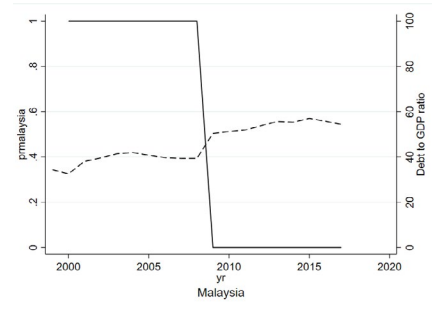


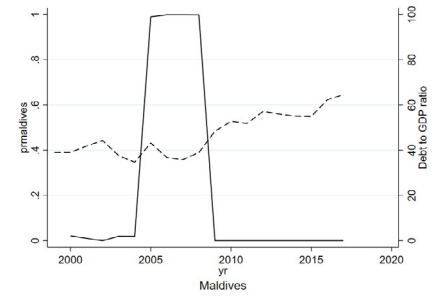
Appendix. Model estimated smoothed probabilities of State 1 (continued)

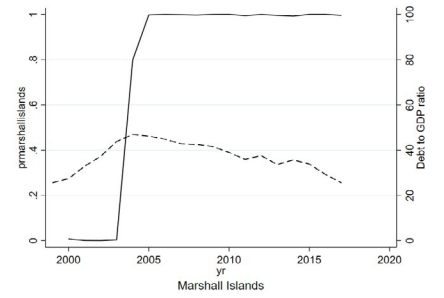


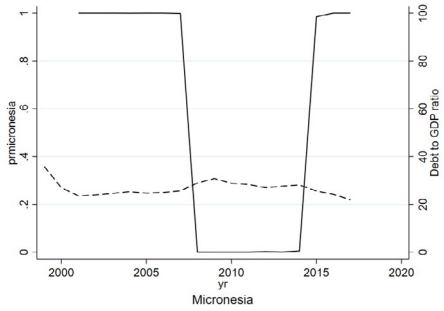


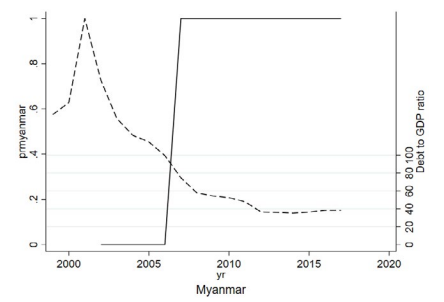




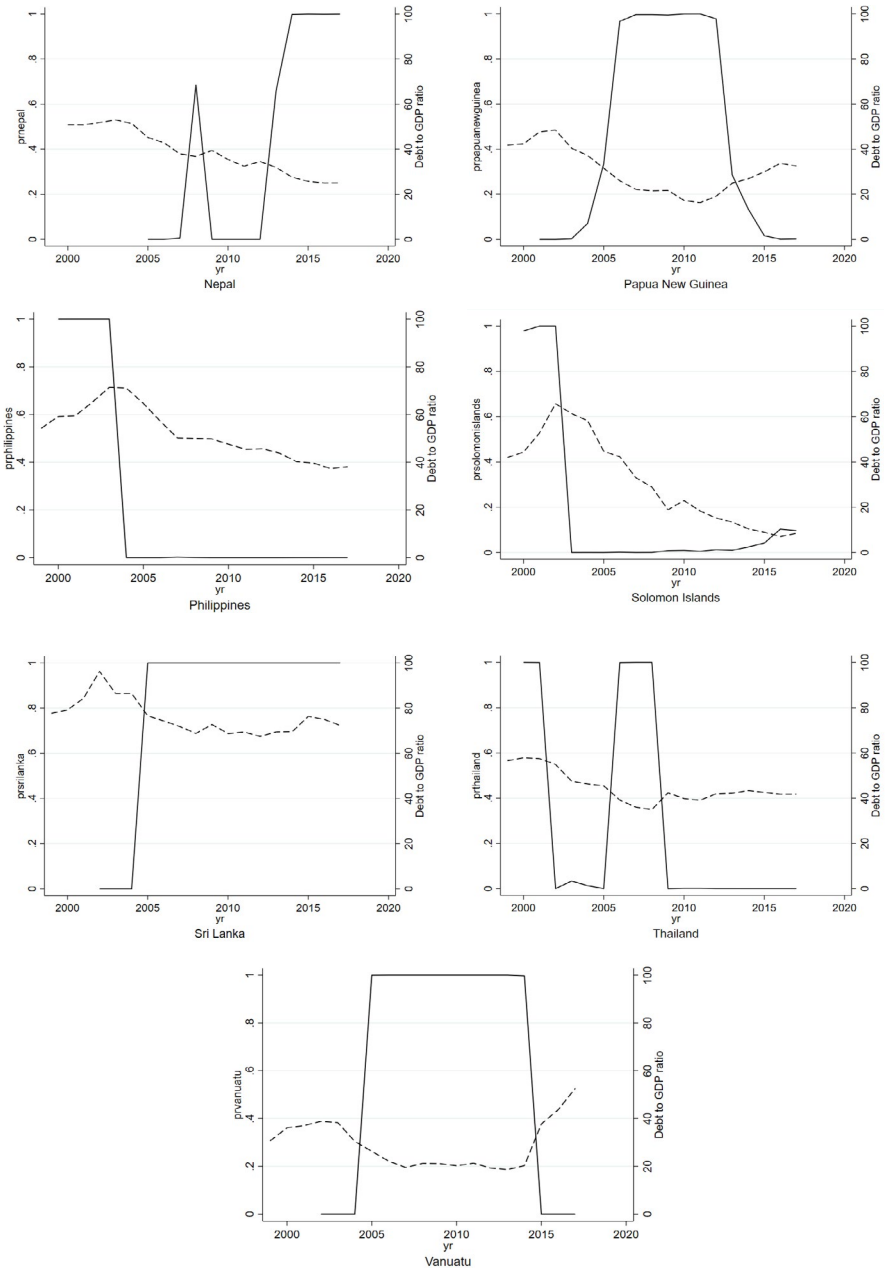








Appendix. Model estimated smoothed probabilities of State 1 (continued)



Notes:
[1] Solid lines are the smoothed probabilities (left axis), dashed lines are the debt-to-GDP ratio (right axis). Since State 2 smoothed probabilities are a mirror image of State 1 in two-state models, these are not shown.
[2] The graphs of Bangladesh and Vietnam are excluded since their smoothed probabilities of State 1 were not obtained.

The Philippine Review of Economics
60(2):104-137. DOI:10.37907/6ERP3202D

The 16th century *Carrera del Pacífico*: its sailor-merchants and their trade goods

Kristyl Obispado*
University of the Philippines

This article focuses on the sailors who served during the initial years of the *Carrera del Pacífico*, one of the globalizing projects of the Spanish Monarchy. In particular, the paper aims to examine the sailors who took advantage of the Pacific trade circuits by actively participating in various income-generating activities created by the *Carrera*. Using the sailors' economic endeavors, especially as sailor-merchants, as a lens can elucidate how early global trade was conducted and demonstrate the dynamics of the early Pacific trade. The paper argues that by seizing the opportunities presented by the *Carrera*, primarily by assuming the dual roles of sailors and merchants, these laborers helped consolidate the Spanish-Pacific region and reshape the consumption pattern of its local population. The sailors engaged in the transportation, sale, and purchase of global commodities during the early modern period, including textiles and chinaware, which catered to the demands of the broader consumer base in Spanish America. The sources draw data from the Royal Treasury of Acapulco registers during its first decade (1590-1600), where 1,574 sailors were identified. It belongs to Archivo General de Indias' *Contaduría* (Account) records, which contain the duties of commodities entering and leaving the port of Acapulco.

JEL classification: F13, B15, B17, N76
Keywords: Pacific trade, sailor-merchants, early globalization, Philippine-Chinese goods

1. Introduction

The term *Carrera del Pacífico*[1] (*Carrera* from here forward) or Pacific trade is used instead of its more popular name, Manila-Acapulco trade, to have a more nuanced understanding of its formal and informal traffics, including the Canton-Manila-Acapulco-Lima route [Bonialian 2012]. The galleons and their incessant navigations attempted to mark and connect the Spanish commercial highways in the West (Spanish America) and East (Asia) Indies. Due to this trade, the massive encounter of people from different continents led to the consolidation of the global spatial network and, to a certain extent, helped change the consumption pattern in Spanish colonies in the Pacific region.

* Address all correspondence to knobispado@up.edu.ph.

[1] The terms *Carrera de las Islas Filipinas*, *Carrera de Nueva España*, and *La Real Armada de Filipinas* also appear in the archival documents.

This study's examination of the pre-1593 period is crucial in understanding the intercolonial commercial interactions before royal policies were created to regulate trade. It should be emphasized that from the time of the *Carrera*'s establishment in 1571 through its first decades, the general nature of the transaction as tax-free endured. The royal decree issued on April 14, 1579 allowed any Philippine ships to trade with New Spain, Peru, Guatemala, and Tierra Firme (part of the present-day Central and South America) without restrictions, a landmark and irreversible act that opened the Spanish-American market—formerly the exclusive domain of the Portabelo-Europe circuit—to Asian products (Schurz [1918]; Borah [1975]; Iwasaki [1992]; Bonialian [2014]). As a result, the first decades, which witnessed the local-global transformation of spaces with the increased movements of ships, people, and goods, significantly influenced its trajectory for centuries to come.[2]

In 1593, even when the Spanish Crown began to implement its trade as a government monopoly, its initial ambivalence towards strict trade restrictions between its colonies resulted in the persistence of triangular trade between Asia, New Spain, and Peru through contraband activities [Borah 1975]. Regarding the royal commerce between Acapulco and Manila, when the Spanish Crown imposed restrictions, reducing the annual trips of 250-300-ton ships from four to three, with the third vessel serving as a backup in emergencies, colonial officials compensated for this by augmenting the ship sizes. Another repeatedly used loophole involved designating King's vessels for non-commercial purposes such as dispatch, exploration, expedition, and supply procurement. Moreover, a notable change occurred with the introduction of higher taxes by 1606. Previously, the tariff on imported goods in Manila had been fixed at three percent. However, it was raised to six percent and remained at this rate until the 1640s. In addition, export goods destined for New Spain faced a two-percent tax alongside a 12-peso ship freightage fee per ton [Qing 2018].

The Crown's gradual regulation of the *Carrera* significantly favored the royal sailors. But even before the 1593 prohibition, it must be emphasized that they had already been benefitting from the trade.[3] Since there was no restriction to the quantity of goods officials and crew could carry, the sailors had been active participants since the very beginning (Álvarez [1993]; Álvarez de Abreu and Yuste [1977 (1736)]; Recopilación [1973]; Obispado [2021]).[4] When the 1593 royal order prohibited residents of New Spain and other parts of Spanish America from engaging in the trade, sailors capitalized on their privilege to transport and traffic commercial goods. With this setting, the Pacific sailors played protagonist

[2] For more details about the early phase of Pacific commercial activities, see Borah [1975] and Iwasaki [1992].
[3] Iacarrino [2011] notes that before the establishment of the *permiso*, Spanish merchants enjoyed a high degree of freedom, allowing them to load whatever quantities of goods they desired onto multiple ships. However, the same royal decree imposed restrictions, limiting the size of the galleons to a mere 300 tons.
[4] It was in 1602 when a decree was first issued regarding cargo distribution in the galleon: among the 300 tons of load distributed in the ship, each Spanish sailor had a right to half a ton while each *grumete*, *a fardo* (1/8 of a ton).

roles as merchants during the early trade. They helped to incorporate global commodities such as textiles and chinaware into the Spanish-American market. It is evident in the archival data, which reveals that throughout the 16th century, ten percent of the sailors took part in the trade route where they served as sailors: Manila-Acapulco, Acapulco-Peru and Central America, and Peru and Central America-Acapulco.

Documented on the *Cuentas de Real Hacienda of Acapulco* (Account of the Royal Treasury of Acapulco), one finds a meticulous account of trade commodities and their *almojarifazgo* or duty collected in Mexico. Across a thousand folios of *Contaduría* number 898 is a trove of information, including details of apportioned spaces, names of the owners or consignors of goods, and the three designated individuals responsible for handling the disposal of commodities in Acapulco. Initially, as mentioned before, the duty was at a rate of 12 pesos per ton and later increased to 45 pesos in 1586. However, by 1591, it was fixed at a ten percent ad valorem rate. The proceeds generated from the *almojarifazgo* were then remitted to Manila as a component of the *situado*.[5]

Given the detailed record of every sailor's consignment and corresponding duty, these sources shed light on the operation of 16th century trade and present some of the earliest inventories of goods transported by sailors, often referred to as cargo manifests. The primary strength of this data lies in its documentation of the actual commodities that were consistently exported, marketed, and consumed across various regions. It reveals a significant shift in trade dynamics, moving away from premium products like spices toward essential products routinely carried by ships. This transformation could signify the diminishing gap between producers and consumers as they became closely interconnected within a global trade network. Furthermore, these records demonstrate the trade of goods with a global reach, encompassing affordable consumer products designed to cater to diverse household needs. These items span a wide range, from clothing and bedding to kitchen utensils, providing information on household consumption patterns in the early modern world.

This study does not consider contraband goods loaded onto ships, which might include luxury items such as spices and jewelry. Notably, during this specific decade, there was a substantial surge in silver importation to China, estimated at 6,000,000 pesos. Apparently, the 500,000-peso limit was exceeded by more than tenfold, indicating that it is far from accurately representing the true extent of the trade volume [Bonialian 2021].

2. Sailors as merchants

The trade activities of the sailors during the initial period of the Pacific trade were intricately tied to their occupational mobility, especially the voyages from

[5] See Luis Alonso's [2019] comprehensive discussion regarding Philippine royal situado.

Manila to Acapulco. However, their responsibilities extended beyond the Manila-Acapulco route and became instrumental in the Spanish empire's military and commercial expeditions in larger Spanish America and the Asia-Pacific regions. Consequently, these areas became an indispensable source of commercial products for the traders, which included the sailors. This section traces the geographic origin, job history, level of participation in trade, and commercial routes followed by these sailors in the 16th century. In the process, it aims to articulate how a particular non-elite group experienced the global commercial changes during the period. For the sources, it draws data from the registers of the Caja Real de Acapulco in its first decade (1590-1600), where a total of 1,574 sailors, comprising *marinero* (sailors), *grumete español* (Spanish apprentices), *grumete indio* (native apprentices), *paje español* (European pages), and *paje indio* (native pages) are identified. The findings of this study indicate that ten percent of these individuals participated in the trading activities of the ships while serving as sailors. These trade routes included Manila-Acapulco, Acapulco-Peru and others, and Peru and others-Acapulco.

2.1. Sailors' origin and mobility

Set in the context of the early global trade, the *Carrera* generated a lot of jobs and pulled towards its port crew members originating from Asia, Europe, and Spanish America. Asians and Spanish Americans constituted about 36 percent of the crew and served primarily as *grumetes indios*. At the same time, nearly 63 percent were Spanish and non-Spanish Europeans. The enlisted sailors of Asian origin came from the Philippines or other territories with commercial relations with the Portuguese in Macao and Spaniards in Manila during the Iberian Union. From the other side of the Pacific, particularly Acapulco and Mexico City, which served as other key recruitment areas for the sailors, came the *indios novohispanos* (Mexican indigenous population) and *mulattos*. The Europeans entered the *Carrera* through the Atlantic Veracruz, from whence they sailed to Acapulco and were hired as sailors, *grumetes españoles*, and *pajes españoles*.

TABLE 1. Geographic origin of sailors, 1590-1600

Sailors	Geographical Origin	Frequency	Percent
Grumete español (Spanish apprentice)	Europe	153	9.72
Grumete indio (Native apprentice)	Asia and America	568	36.09
Marinero (sailors)	Mostly Europe	825	52.41
Paje español (European page)	Europe	26	1.65
Paje indio (Native page)	Asia and America	2	0.13

Source: AGI [1591-1603].

The geographic origin and religion of sailors determined their rank as *marineros*, *grumetes*, or *pajes*, which would affect their salary and other remunerations. However, examining their job history suggests that their post as sailors was highly flexible. Table 2 shows that among the identified 1,574 sailors (*marineros*, *grumetes españoles*, *grumetes indios*, *pajes españoles*, and *pajes indios*), a total of 162 were registered to have served simultaneously or alternately in other positions aboard the ship and at ports between 1590 and 1600.[6] The frequency is the number of times any of these 162 sailors assumed other positions. For example, one or any of them took the position of a shipmaster five times.

TABLE 2. Sailors' mobility across ranks, 1590-1600

Position	Annual salary in *pesos de oro comun*	Frequency	Percent
shipmaster	200	5	1.32
boatswain		1	0.26
assistant pilot	150	3	0.79
assistant to chief pilot	300	1	0.26
water bailiff	200	4	1.05
captain	180-200	5	1.32
gunner	120-200	34	8.95
barber and surgeon	150-250	1	0.26
plumber	83p 2t 8g	1	0.26
diver	300	2	0.53
caulker	300	11	2.89
captain and master	300	1	0.26
carpenter	300	6	1.58
native carpenter	100-250	7	1.84
chief gunner	250-300	6	1.58
shipmaster	250-300	11	2.89
dispenser	200	13	3.42
scribe	180	1	0.26
grumete español	100	40	10.53
grumete indio	50	21	5.53
grumete indio de chinchorro	50	1	0.26
port's grumete indio	50	3	0.79
steward	200	16	4.21
artilleryman	83p 2t 8g	4	1.05

[6] The author utilized the appendix attached in Obispado [2021] to reconstruct the number of sailors who served in various positions in the Pacific trade between 1590 and 1600.

TABLE 2. Sailors' mobility across ranks, 1590-1600 (continued)

Position	Annual salary in pesos de oro comun	Frequency	Percent
maestre	300	4	1.05
marinero	150	147	38.68
port's marinero	120-150	4	1.05
paje español	80	2	0.53
paje indio	40	0	0
pilot	600	3	0.79
chief pilot	600	1	0.26
cooper	300	2	0.53
turner	83p 2t 8g	1	0.26
port's royal employees		18	4.73
Total		**380**	**100**

Source: AGI [1591-1603].

It means that even a *grumete indio* whose annual salary was 50 pesos could still take other jobs as captain, carpenter, and caulker and receive up to 250 pesos salary. Shifting between two or three jobs or being employed simultaneously for two to three positions also implies a greater chance of engaging in trade and increasing one's income. While waiting for their next voyage, they would take other jobs at the port. Finally, they alternately participated in different *Carrera* trade circuits: the Manila-Acapulco, Acapulco-Peru, and Peru-Acapulco. Between 1591 and 1600, 59 ships plied between the ports of Manila and Acapulco, which explained the demand for sailors [AGI 1591-1603]. The mobility observed among crew members hints at a potential crew shortage and budget constraints within the Treasury of Acapulco and highlights the sailors' ways to improve earnings. Their effort to augment their salary becomes particularly evident as they capitalized on the opportunities that the *Carrera* presented. The sailors had access to information about various goods and producers, fostered crucial contacts within their workplaces, and engaged in commercial interactions as they embarked on intercontinental journeys. Moreover, they were allowed to bring along slaves, a trunk for personal belongings, and extra luggage for their merchandise aboard during their voyages.

2.2. The sailor-merchants

Due to their occupation, sailors frequently found themselves in bustling commercial centers where goods circulated. They leveraged their privileges to generate income either by selling their allocation of a galleon's cargo space to non-Manila residents, by being owners or consignors, or by acting as consignees, earning a commission from those who were not eligible to engage in trade

(Ortigosa and Lee [2018]; Goode [2017]). During this period, some sailors temporarily left their maritime posts to concentrate on trading. Subsequently, several of them returned to their roles as royal sailors, while others did not. For example, Francisco Diaz, who served as a sailor in 1591, became a *consignatario* (consignee) in 1593 before resuming his service as a *Carrera* sailor in 1594, a post he maintained into the 17th century. Another case is that of Benito Juarez, who served as a sailor in 1591. However, by 1594, he repeatedly emerged as a consignee in Contaduría records. Francisco de Bolaños and Bernabe de Vera shared a similar trajectory. The former worked as a sailor from 1590 to 1591 but transitioned into the role of a trader in 1592. In contrast, the latter juggled his work as a sailor and trading activities in 1592. However, by 1594, Bernabe de Vera had left working as a sailor and dedicated full time to commerce, as his name no longer appeared as a sailor but kept surfacing as a trader until 1598 [AGI 1591-1603].

Another observation is the active involvement of *marineros* and *indios grumetes* in the commodity trade. In 1592, several *indios grumetes* jointly traded goods valued at 350 pesos while traveling from Manila to Acapulco [AGI 1591-1603]. However, individuals in higher positions often made more significant investments. Consider the case of Francisco de Landia. He began as a sailor on February 28, 1591 and advanced to becoming sailor-boatswain the following year, on February 10, 1592. In the same year, he began participating in trade along the Manila-Acapulco route. He continued working as a sailor-boatswain until May 19, 1594, when he became a pilot from May 20 until July 1, 1595. Simultaneously, he engaged in trade with higher capital. His diverse positions included serving as a gunner, *marinero*, *grumete español*, shipmaster, and steward, which continued until 1622 [AGI 1591-1603].

Similarly, a trader named Antonio Rodriguez started his career as a pilot-merchant in the Manila-Acapulco route on July 8, 1592. On April 9, 1593, he remained as a pilot while trading in the Acapulco-Peru route. In 1594, he continued his service on dispatch ships involved in the galleon trade, with his trade goods on board. Toward the latter part of 1594, he again shifted his focus to the Acapulco-Peru route. Upon his return to Acapulco, he brought products from Peru and paid the five percent *almojarifazgo*. In 1595, he resumed his duties as a pilot on royal ships traveling from Acapulco to Peru. Concurrently, he purchased goods in Acapulco [AGI 1591-1603].

The last example is Gaspar Ramirez, who served as a steward for the Acapulco-Manila-Acapulco route from December 1, 1590 to June 12, 1591. Although it is not explicitly documented, he must have returned to Manila toward the latter part of the year, given that on July 8, 1592, he was en route from Manila to Acapulco, serving as a sailor. At the same time, he took advantage of his job by bringing trade goods. He continued in the capacity of a steward until January 26, 1593, after which he became a dispenser from January 27, 1593 to December 31, 1594.

He also registered trade goods within this period [AGI 1591-1603]. This pattern of dual involvement in trade and official roles extended throughout the service of the officials and crew. However, as the 17th century progressed, the documentation of sailors' commercial activities gradually faded. This shift might have been a strategic move to obscure contraband trade practices. As a result, the register of *almojarifazgo* (also called *flete de ropa*) ceased to provide specific information about the owners and brokers of the trade goods and only recorded the amounts they paid for the galleon commodities they loaded onto the ships.

If we survey the commodities they owned, most were textiles (especially silk and cotton), chinaware, furnishing, and other stuff, which came from their workplaces, such as Acapulco-Peru, Peru-Acapulco, and Manila-Acapulco trade networks [AGI 1591-1603]:

TABLE 3. Number of sailor-merchants from 1590-1600

	1591	1592	1593	1594	1595	1599	Total
Acapulco-Peru		1	10		1		12
Peru-Acapulco			3	2			5
Manila-Acapulco	1	76		36		9	122
Total	**1**	**77**	**13**	**38**	**1**	**9**	**139**

Source: AGI (Archivo General de Indias). 1591-1603. Caja de Acapulco, Cuentas de Real Hacienda [Treasury of Acapulco, Accounts of the Royal Treasury]. Contaduría 898-901.

The limited presence of sailor-merchants on the Acapulco-Peru and Peru-Acapulco routes can be attributed to a series of trade restrictions that came into effect, including an early prohibition on trade between Peru and the Philippines dating back to 1581. Subsequently, in 1591, another ban was imposed, which forbade inter-American commercial interactions from Callao to Tierra Firme and Guatemala. By 1604, all trade between the two viceroyalties of Mexico and Peru was completely prohibited [Iaccarino 2011:116]. Thus, a striking contrast emerges when we examine the situation in the years preceding these bans. For instance, all the sailors and crew members aboard the ship *Santa Maria de la Cinta*, which arrived in Peru in 1581, were actively involved in trading activities as owners and consignees.[7]

The peak in the number of sailor-merchants recorded in 1592 (Table 3) offers a significant glimpse into the actual engagement of the *Carrera* crew in trade activities. During the initial years of the *Carrera*, commercial ties between the Philippines and China remained unrestricted. However, changes began to take shape in 1586 when Spanish authorities petitioned the Council of the Indies for a shift in how purchases from Chinese merchants were conducted. This led to the request that individual arrangements be replaced by a wholesale agreement endorsed by a royal decree in 1589 [Iaccarino 2011]. This transition was

[7] For the details of the trade goods aboard the ship *Santa Maria de la Cinta*, see Obispado [2023b].

ultimately enforced in 1593, along with new regulations governing the trade, which prohibited voyages to China for trade purposes and cessation of importing goods through Chinese vessels specifically consigned to individual Spaniards in Manila. Instead, the only permissible method for procuring Chinese imports became the *pancada* system, a wholesale bargaining mechanism. It was designed to regulate the volume of goods acquired annually for each galleon trip [Schurz 1939]. This shift in trade practices had a notable impact, as substantiated by the significantly more extended list of shippers in the 16th century when contrasted with the 18th century [Schurz 1939].

Post-1593, the merchants started resorting to contraband trade, looking for the loopholes of the royal decrees. Schurz details these practices, noting that merchants would set up specific shipments of goods from China at prearranged rates. The consignee would receive these consignments before the junk reached Manila Bay, or some merchants would buy goods ahead of the official appraisal, which involved a comprehensive examination of goods or samples with intricate classification and grading. Some opted to invest after the *pancada* had concluded, storing silk in private warehouses to await the sailing of the following year's galleon. Others chose to invest after the *pancada* had finished, holding silk privately to be dispatched on the following year's galleon [Schurz 1939].

3. Expanding market for Asian goods

As seen in the previous section, the workplace mobility of the sailors in the 16th century *Carrera*, in conjunction with their roles as traders and intermediaries, provides insights into Pacific commercial exchanges. In this time of early market expansion, the sailors were strategically positioned to influence the changing consumption pattern, especially in Spanish America. The surge in consumer demand can be attributed to increased access of the lower classes to silver and silk given the availability and cheapness of goods. By tracing the involvement of sailors in the Pacific trade networks, we can also gain a deeper understanding of how products underwent adaptation to align with the ever-evolving preferences of consumers as they became more intricately connected to the global trade system.

The variety and quantity of goods carried by the ship crew depended on the route of their journey [Obispado 2023a]. However, the context of the period in which the trade was conducted also played a crucial role. During the initial decade of the *Carrera* (1570s-1580s), coveted spices, such as cloves, pepper, sesame, nutmeg, and cinnamon, were exported by the officials and crew for up to more than 20,000 pounds. But between the late 16th and early 17th centuries, a few pounds of cinnamon became the only spice appearing among the items they transported. This period's commodity profile became dominated by cheaper raw, intermediate, and finished Philippine-Chinese goods. This section analyzes the context of the period that created the demand for Asian fabrics, earthenware, and other products.

It is an attempt to unbox and examine the tightly crated galleon cargo, an action that was inconceivable at any port, particularly from 1593 onwards.

3.1. The pre- and post-1593 period

From the economic angle, the *Carrera* became instrumental in the changing global landscape of production, consumption, and circulation of goods, gearing toward large-scale production. One of the earliest indicators of this new paradigm was the emerging and overlapping intercontinental commercial circuits, which encompassed geographical, political, commercial, and cultural dimensions. Moreover, the locations where this transformation was unfolding—ports, towns, urban centers, and hinterlands—relied on various network systems constantly expanding, contracting, and creating formal and informal venues for commercial transactions.

Interconnected global events marked the 16th century. Synchronous with the discovery of silver in Spanish America was the shift towards a silver-based economy in China. The conjunction of the low production cost of silver in America and its high demand in China gave birth to global trade [Flynn and Giraldez 1995]. It is crucial to underscore earlier events that propelled these global transformations: the Ming government's choice to open a part of the Fujian region for trade in the 1560s, followed by the establishment of the Spanish colonial outpost of Manila in the Philippines during the 1570s [Qing 2018]. The transfer of the Philippine capital from Cebu to Manila highlighted the government's intent to prioritize commercial ties with China, signifying a shift away from the trade in spices and toward silk. Because of its strategic location and its integration into preexisting Asian trade networks, Manila evolved into a commercial center. It became a pivotal conduit for exchanging luxury items catering to the elite while streamlining the regular exchange of bulk commodities between Asia and Spanish America [Gerritsen and Riello 2015]. Another significant episode was the dynastic union of Spain and Portugal, aligning with increased economic activity in Asia. As Tremml [2012] and Pinto [2014] posited, the Iberian Union amplified economic interactions between their Asian territories in the South China Sea, notably Manila and Macao, which further expanded to Melaka, India, and Japan. As a result of the abovementioned 16th century events, the Pacific World, the New World, and the Old World had indeed become permanently linked [Flynn and Giraldez 2008].

The second indicator was the increasing enforcement of the *almojarifazgo*. This term, derived from the Arabic word “al-musrif”, referred to the royal tariffs established for all Spanish colonies. In the Philippines, it was first introduced in the 1580s by Governor Gonzalo Ronquillo de Peñalosa when a three-percent tax was proposed on import and export goods in Manila and Acapulco as well as the 12-peso ship freightage for each ton on ships heading to New Spain. Qing [2018] considers 1591 to be the commencement of large-scale Philippine-Sino trade. This assertion is grounded in Juan Gil's comprehensive record of Chinese captains’

names and the customs duties paid upon their ships' arrival in Manila. This year, twenty-one Chinese ships registered and paid the three percent *almojarifazgo* in Manila. It translated to a substantial tax revenue of 16,829 pesos, five *tomíns*, and four *granos*, levied against the imported goods valued at 560,967 pesos.[8] The evident expansion of trade compelled the Spanish crown to regulate it through strict implementation of taxation on the galleon goods. Eventually, the acceptable amount of merchandise *naos* (galleons) could transport from the Philippines to New Spain was limited to 250,000 pesos, while the monetary earnings should not exceed 500,000. But even with this limitation, the Chinese continued sending products to Manila that could reach up to several million pesos and stayed this way until 1615 [Qing 2018]. Given the 250,000-limit to the merchandise permitted to be brought to Acapulco, most Chinese goods could have been sent as contraband.

For the last indicator, the trade landscape from the 1570s, dominated by exotic goods such as spices, began to disappear. In contrast, what used to be considered a luxurious item, such as silk, became accessible to everyone, which somehow leveled the social rank among people. The Spanish Crown began implementing sumptuary laws in its various colonies to curb this. For example, in the Philippines, Governor Dasmariñas passed an ordinance as early as 1591 prohibiting Filipinos from wearing Chinese clothing [Iaccarino 2011].

3.2. Shift to Philippine-Chinese goods

One tangible outcome of the 16th century events was replacing Andalusian goods in Spanish America with Philippine-Chinese goods when the *Carrera* began. This shift could be elucidated by examining the composition of the Spanish trade monopoly. They primarily comprised raw or manufactured textiles crafted from silk, woolen, or cotton. While there was a sustained and constant demand for fabrics, the faltering Spanish textile industry needed help to meet this demand. Moreover, its products were considerably more expensive than those of its Mexican counterparts.

In Mexico, the local authorities had already developed their silk industry as early as 1522, cultivating mulberries and producing silk textiles [Iaccarino 2011]. A survey of textile goods that arrived in Puebla between 1549 and 1562, conducted by Boyd-Bowman [1973], highlighted substantial price disparities. For instance, Mexican silk, primarily sourced from Granada, was retailed for less than a third of the cost of Spanish imports. Similarly, locally produced taffeta was priced 25 percent lower than the average imported equivalent.

Meanwhile, as the galleon system evolved into an efficient means of cost-effective and systematic distribution, it was necessary to introduce novel and practical goods to broader markets. It started in 1573 when two ships brought 712 bolts of Chinese silk to Acapulco, 22,300 pieces of fine china gilt, and other

[8] Eight silver *reales* or *tomines* made up a silver peso; while each *tomín* could be divided into twelve *granos*.

porcelain wares [Villiers 1980], eventually dominating the American market. Even the silk produced in Mexico could not match the quality of China's silk, which was available at nearly a third of the price of those made in Spain [Iaccarino 2011]. Leveraging the strategic nexus between Manila and Acapulco, as well as the burgeoning intra-Asian trade extending to Central and South America, the Philippines and China emerged as pivotal hubs for the production of galleon goods, particularly textiles, in the 16th century.

From the Philippines, the survey of goods showed medrinaque and cotton as the most common products. Unsurprisingly, fabrics became core components of the early tribute payment system in the colony. For instance, in Manila, a standard non-monetary tribute included 185 pounds of rice, one cotton cloth that measured two to 2.6 yards long and one to 1.2 yards wide, one mae of gold, and one chicken (Obispado [2023b]; Hidalgo [1995]; Alonso [2019]). In the Visayas, *guinaras*, a cloth measuring 15 feet by 20 inches, became synonymous with *tributo* or tribute since they were used to pay tax [Castro 2018]. The fabric measurements stipulated for tribute are particularly noteworthy, suggesting the nature of the Philippines' principal exports destined for Mexico. It should be emphasized that from the precolonial period, the local populace, particularly women, had been actively involved in community and household economic tasks like weaving for internal consumption and fulfilling their traditional tribute obligations. With the rule of the Spanish Crown, they were continuously tapped to meet the growing demands for textiles in New Spain and other parts of Spanish America (Obispado [2023b]; Diaz-Trechuelo [1964]). In producing these handwoven fabrics, numerous processes, from stripping and dyeing abaca to spinning and weaving, were carried out in different households, especially by women.

The galleon system attracted manufacturers in China, who found a ready market for their goods in the *Carrera* merchants exporting to Mexico, Peru, and Spain [Trusted 2013:448]. The initial tax-free policy extended by the authorities was a significant incentive for exporters, most of whom were Chinese. Even when Governor Dasmariñas (1590-1593) set a six percent tariff on all Chinese goods, it did not deter the influx of Chinese traders. In fact, during the years spanning from 1581 to 1590, an average of 102 Chinese vessels arrived in Manila, and this number increased to 119 during the period from 1591 to 1600 [Iaccarino 2011]. These junks provided an assortment of silks, from raw materials to elaborately woven textiles such as velvets, damasks, and taffetas, each demonstrating a variety of weaves, textures, and shades. Additionally, there was a rich display of linens, cotton shawls, items adorned with gold, and intricate embroideries [Schurz 1939].

In addition to manufacturing hubs in China, workshops in Manila were another source of Chinese products. Caniquís and Sangley's thin mantas, which could have been produced in these workshops, were priced as affordably as native-made mantas, both carrying a two-tomín-price (Obispado [2023b]; AGI [1592]). As Bishop Salazar noted in his report to the king dated June 24, 1590,

some articles were crafted by Chinese artisans within Manila's Parián, and they did so "quickly and with better finish than in China" [Iaccarino 2011]. Workshops within this area were established starting in the late 16th century, producing both standard and luxury materials, including items made from ivory.

Trusted [2013] pointed out that the manufactured articles offer insights into their operation, hinting at broader workshop practices. The diversity of ivories discovered suggests that various versions or possibly duplicates of the same subject and composition were produced in significant quantities. Hence, it is likely that there was an existing production line, with the repetitions indicating not only an efficient method of production but also a consistent demand for popular sacred images.

In the context of the galleon trade in the 16th century, it becomes apparent that there was a presence of household or local manufacturing along with workshops in Manila, other capital cities, and central hubs in China, each engaged in the production of goods at various stages, encompassing raw materials, intermediate products, and finished goods.

4. Unboxing the shipped goods

Ortigosa and Lee [2018] underscore the complex and laborious nature of documenting merchandise for transport. Multiple variables are at play, such as the prevailing market conditions, the cargo capacity of the ship, and the strategic allocation of goods to optimize efficiency and profitability. This meticulous approach ensured that each voyage was tailored to meet the specific demands of the time while satisfying the interests of the shippers and the broader trade network. For the standard process, every package had to be sealed, labeled, and registered for taxation in the *libros de contabilidad* (manifest of goods). Upon arrival in Acapulco, the port officials headed by the *castellano* (port governor) would board and inspect the vessel, collect the official manifest of goods, dispatch them to the royal treasury, and then facilitate the unloading of the cargo to be transferred to and kept in the warehouse until the Acapulco fair (Goode [2017]; McCarthy [1993]. To carry the merchandise ashore, they used *chata*, a flat-bottomed boat that the *cagayanes* (shipyard workers) usually fabricated for over 20 days.[9] The officials paid the sailors four pesos daily to empty the ship's cargoes and appointed guards to monitor them closely. The activity usually lasted two weeks; however, other challenges, such as the arrival of vessels carrying sick crew members or the appearance of Spanish enemies, could prolong it [AGI 1606-1615].

For the taxation, the shippers had to provide a detailed account of their consignments and a declaration specifying each chest's volume and contents. In addition, they were under the obligation to formally swear, signifying their

[9] They spent four days going to the mountains to cut and bring small knee, futtock timbers, and other woods; another 16 days to assemble it and its oars, from which they received a fee of five pesos per day. [AGI 1606-1615].

solemn acceptance of the stated information. This oath affirmed that the goods corresponded precisely to the provided description and that the shipper was the exclusive and original consignor of the shipment. Following the board of appraisal certification, the invoices were submitted to the *Contaduría*, which were meticulously copied into the galleon register. Then, the duties were assessed based on the declared value provided under oath [Schurz 1939]. The *almojarifazgo*, or tax owed to the King, depended on the origin of the vessels: ten percent if they came from the Philippines and five percent from Peru and other American ports. These taxes were sent back to Manila as part of the Philippine situado (AGI [1591-1603]; Schurz [1939]; Alonso [2009]).

However, as noted by several authors, it was almost impossible to appraise the ship's overall cargo without examining the goods. Depending on the container, one ton could be filled with approximately eight *fardos*, six *cajas*, or 12 wax cakes. Containers did not adhere to standard form and instead varied in size, weight, or capacity, as follows: *fardo* (bale), *fardo grande* (large bale), *fardo chico* (small bale), *fardillo* (small bale), *fardillo chico* (small fardillo), *cajuela* (small box), *cajon* (crate), *caja* (box), *cajoncillo* (small crate), *petaca* (chest), *balsa* (raft), and *lío* (bundle). Sometimes, they even utilized *escritorio* (desk) to pack the goods (Ortigosa and Lee [2018]; AGI [1591-1603]). Hence, opening the boxes of goods was the only method to ascertain the invoices' value. Due to the limitations in unlocking or untying the containers, examining the traded goods and confirming whether what was recorded in the manifest matched the merchandise's actual type, quantity, and value was nearly impossible. According to Ortigoza and Lee [2018]:

> The vast variety of merchandise and the multitude of different units of measurement assigned, sometimes even for the same type of merchandise, along with the varying value of that unit depending on the time and port of registration, and the fact that the same product could be noted in different measures of capacity, weight, or length, all contribute to the complexity. Different records for the same type of merchandise introduced into the ship in various measurement forms, by a single merchant or multiple owners or representatives, on different days, and representing different amounts, serve as a brief example of this complication.

Before establishing the 1593 monopoly, three to four ships weighing 250-300 tons made yearly trips for the royal trade between Manila and Acapulco. For instance, in November 1592, three galleons, namely *Santiago*, *Nuestra Señora del Rosario*, and *San Pedro*, arrived in Acapulco. Since port authorities examined and appraised the goods brought by their crew, the ten percent tax they paid was as follows: *Santiago*, 2,104 pesos, three granos; *Nuestra Señora del Rosario*, 583 pesos, three tomines, three granos; and *San Pedro*: 3,078 pesos, one tomín. These taxes would be charged to 76 crew and an unknown number of *indios grumetes* who loaded goods either on their account or as consignees of other merchants.

The ten-percent tax for the merchandise they brought amounted to 5,760 pesos, four tomines, six granos de oro común, or almost 60,000 total value of goods, equivalent to nearly 25 percent of the permissible total value of goods.

4.1. Sailors' inventory of goods

An inventory of trade commodities owned by the galleon crew is shown in the Appendix encompassing a diverse spectrum of 15 distinct categories: food and spices, lead, personal items, home essentials, live animals, slaves, raw materials, woven fabrics, textile accessories, head coverings, personal cloths, garments, footwear, beddings and curtains, and others. Furthermore, this extensive array could be further divided into a staggering 229 varieties, contingent on specific attributes such as their place of origin (Philippines, China, Japan), size (small to medium-sized), dimension (six to 11 yards), weight (lightweight), thickness (thick, thin), color (white, painted, dyed), quality (ordinary, unrefined), texture (coarse, fine), authenticity (imitation, damask-like), material or fiber composition (cotton, silk, linen), types of woven fabrics (grogram, damask, satin), accompanying accessories and decorative finishes (braided, embroidered, brocaded, striped), and processing stage (raw, intermediate, and finished).

The majority of these products came from Asia. Philippine goods included cotton yarns, medrinaque, Moro and Ilocos woven textiles, and other cotton cloths, comprising 50 percent of the total exports.[10] The indigenous population used them, especially plain white cotton or abaca cloth, to make pants, loose shirts, waistcoats, or women's garments. As I have discussed in another article, among the Asian woven fabrics, those from the Philippine islands, which composed 50 percent of the galleon trade goods in the 16th century, were the cheapest, with the price ranging from two to four tomines depending on the types, patterns, and technique used to manufacture the textiles [Obispado 2023b]. Other Philippine manufactured goods that became regular export throughout the galleon trade were Cebu gauze; Ilocos sailcloths, petticoats, and hammocks; Lubang and Ilocos linen sheets; Manila cotton stockings; and tablecloths, bed canopies, and coverlets from other provinces [Schurz 1939].

Chinese mantas were also affordable, albeit costlier than those from the Philippines. Most of their products were made from lower-quality materials and had undergone rapid production. They were described as *ordinaria* (ordinary), *llano* (plain), *común* (common), *basta* (coarse), *cruda* (unrefined), *pequeña* (small), *angosta* (lightweight), loosely woven, and imitation, priced for as low as six *tomines*. Examples are Cantonese and Chincheo's ordinary and unrefined woven textiles. According to Iaccarino [2011], these goods were made in Jiangnan factories and sent through Cantonese and Fujianese maritime trade networks. The latter, known for its mulberry and porcelain production, became

[10] In another article the focus of discussion is the Philippine local produce in the sixteenth-century galleon trade. See Obispado [2023b].

hubs where shipowners could set up shipping businesses and export goods from Fujian and other provinces. Later on, with the endorsement of the Ming provincial authorities, these items, mainly the Chinese silks, were transported by Fujian merchants to Manila. One of the most common items inside the sailors' boxes was labeled based on the ports where the textiles originated, such as Chincheo and Lanquin. From Lanquin came blankets, plain satin, silk cloth with satin stripes, linen, and cotton (Yuste [1984]; Castro [2018]).

The most expensive components on the list of woven Chinese textiles underwent a more intricate process: for example, damask, grogram, taffeta, satin, and fabrics that were embroidered, adorned, or finely colored. Considered premium goods were *raso negro de Lanquin* (black satin from Lanquin), followed by *raso negro de Canton* (black satin from Canton), priced at 13 and 11 pesos *de oro común*, respectively. Also costly, which sold for five to six pesos, were *telilla de oro y seda* (fabrics woven from gold or silk) and *raso de Lanquín de color fino o negro* (Lanquin's fine-colored satin), both sold at six pesos; *damasco ordinario o común de 11 varas* (ordinary damask that measured 11 yards), black or colored satin (usually from Canton), and *tafetán de Lanquín* (Lanquin taffeta), each for five pesos [AGI 1592].

Regarding the processing stage, the more economical raw materials like cotton or silk yarns were prevalent. They were available in loose and twisted forms, typically sold by weight or quantity. Notably, the twisted variety commanded a higher price than the floss thread [AGI 1592]. However, it is worth highlighting that the list was dominated mainly by intermediate or semi-finished products, with mantas or woven fabrics taking the lead. The term manta encompasses a wide range of ordinary cloth types produced in various regions. This can include coarse cotton or linen used as bed covers. Furthermore, as indicated by Castro [2018], it is a general term for everyday fabric manufactured in various locations, including the Philippines. The preference for these cloths in Mexico, especially the unbleached cotton fabric, stemmed from its appeal to the lower-end market, where it found significant consumer demand.

These textiles manufactured in large volumes came in standard dimensions, varying from 5.5 to 11 yards, ready to be adjusted based on the clothing needs and specifications of the target consumers. Standard measurements were employed for various woven fabrics: 7.33 yards for general use, 6.87-7.33 yards for ordinary white or colored taffeta, ten yards for damasks, and ten to 11 yards for Chincheo fabrics. These dimensions served practical purposes. For instance, taffeta, measuring 6.87 yards, was used in crafting veils. In the Philippines, the 7.33-yard fabric served as the customary length for *lompot*, utilized by indigenous communities for various purposes such as gauzy canvas-like materials, blankets, cotton fabric for shirts, tablecloths, and sheets [Castro 2018]. Establishing standardized dimensions and colors for these intermediate goods was pivotal in facilitating a more extensive scale of intercontinental trade between Asia and the Spanish Pacific.

4.2. Finished goods

Many goods, produced in large quantities, were intended for everyday use, spanning various categories, including clothing for the body, bedroom essentials, and kitchen items. Ready-made clothing items were available, such as the *camisa de manta llana* (plain cotton shirt) priced at seven tomines, the *jubón de lienzo sin mangas* (sleeveless linen jacket) at one peso, the *ropilla de raso negro* (black satin and double-sleeved short jacket) at one peso, the *jubón de tela* (linen jerkin) at one peso and two tomines, and a pair of satin sleeves at one peso and two tomines. Breeches, known as *calzones*, came in two varieties: ordinary white ones made of manta, sold for three tomines, and higher-priced black satin breeches, which were priced at three pesos. Stockings also had their distinctions, with linen ones costing one tomín, cotton ones at two tomines, and silk stockings considered luxury items priced at two pesos per pair. Additionally, a selection of shoes and slippers was available [AGI 1592]. As can be seen, the most abundant and affordable clothing items were made from mantas, cotton cloth, or coarse fabric. However, there were also numerous satin and silk choices catering specifically to the high-end market.

There was also an influx of clothing articles available during this time, especially those typically made of manta. These items included painted *pañuelo* (kerchiefs), *pañuelo de nariz* (handkerchiefs), *paño de mano* (handcloths), and *pañuelo de rostro* (sweat cloths), with prices ranging from one to two tomines. *Mantillas* (painted scarves) were sold for two tomines, while the embroidered versions commanded a higher price of two pesos. As for *toca* or *toquilla* (headcloths), the plain and unrefined pieces (intermediate goods), measuring approximately six and a half yards, were priced at two tomines and six granos, regardless of whether they were colored or white. *Toca de espumilla* (loosely woven headcloth) was valued at two tomines and six granos. Ordinary *toca de red* (hairnet) had a price of three tomines, while *toca de puntilla* (those with narrow lace edging) was priced at one peso. [AGI 1592].

Affordable and expensive bedding and curtains were also standard among sailors' merchandise. *Almohada de manta* (pillow covers), *sabana de manta* (bedsheets), and *sobrecama de manta* (bedspreads) made of manta were worth two tomines, four tomines, and two pesos, respectively. A bedspread made of *damasquillo* and damask or embroidered could cost between four and six pesos per piece, while embroidered blankets were priced at seven. For *pabellón* (curtains), the ordinary made of manta cost three pesos, and the taffeta was sold at 15 pesos.

Another category of finished goods was home essentials, including crockery. Although there was a presence of fine Macanese plates and bowls or gilded chinas, what was found inside the boxes were mostly common, coarse, and undecorated earthenware, plates, bowls, and saucers. Their lack of elaborate design and uniform price of six tomines per dozen suggest the presence of general consumers

who owned silver to purchase them for everyday use. The use of *cuchara de concha* (shell-like shaped spoon) and *cuchara de nacar* (made of mother-of-pearl) were also popularized in New Spain with their price of two tomines per dozen.

In elite households in Spanish America, additional home essentials such as gilded writing desks, golden escritoires, and various pieces imported from China or Japan were standard. The official Japan-Philippine commercial relationship was only established between 1598 and 1613, under the leadership of Tokugawa Ieyasu [Borao 2005]. However, even before this time, Japanese commodities had already found their way to Manila. Their typical high-in-value items included *escritorio* (desk), *baúl* (trunk), *escritorillo* (small desk), *bufete* (desk with cabinets), and various boxes containing Japanese mirrors and folding screens [Zapatero 2012]. These items were likely commonplace in the daily lives of the elite residents, contributing to the opulence and luxury of their homes.

Personal items include different types of combs (made of wood or tortoiseshell), rosaries, fans, and all sorts of boxes (small, golden, and needle boxes). It is not mentioned of which material the rosaries were made, but in comparison with other holy objects made of precious material, such as large-scale ivory figures destined for aristocratic collections in Spain and Europe, rosaries found aboard the sailors' trade goods also complied with its role as a religious item in the spread of Christianity but were more portable and less costly [Trusted 2013]. All these items clothed a Catholic woman with a veil, comb, fan, and rosary, all cheap and ordinary.

Finally, apart from the investments contained in their chests, officials and crew were also allowed to bring slaves, who were sold in Acapulco for 120 to 150 pesos [AGI 1592]. The ship's officials and crew members had dual roles—they could either act as slave traders or care for the slaves during the journey from Manila to Acapulco. In the latter scenario, the slave owners in Manila hired them to oversee the well-being of the slaves, providing them with food, drink, and accommodations throughout the voyage. Following the journey, it was their duty to sell the slave to the highest bidder in Acapulco and testify to the sale. In return for their services, they received a commission amounting to one-third of the selling price of the slave. Still, it was not until the beginning of the seventeenth century that the cost of slaves, which nearly doubled, ranging between 200 and 400 pesos or even more, was contingent upon the slave's place of origin. Those categorized as "indios chinos" were the least expensive, while black slaves commanded the highest prices (AGI [1604-1632]; Oropeza [2007]; Seijas [2014]).

5. Final analysis

The institutionalization of the *Carrera de Pacífico* entailed a consistent supply of labor to carry out its operations: to facilitate the Trans-Pacific exchange, to deliver subsidies, to transport officials, missionaries, and soldiers to its overseas territories, and to access non-Spanish domains. The participation of sailors in these activities became indispensable, but what has been overlooked in the existing

studies is how they responded to the early global trade. The massive encounter of these people from different continents led to two things: the consolidation of the spatial interconnection of the globe and, to a certain extent, it helped change the pattern of consumption, especially in the colonies in Spanish America. Their geographical mobility as sailors spanned the Pacific Novohispano-Philippine channel and the Mar del Sur that connected the Novohispano with Central and South America. Their protagonist roles as sailors and merchants helped to incorporate global commodities such as textiles and chinaware into the American market. In the end, this paper highlights how the sailors maintained the trade routes and took advantage of these circuits by actively participating in various income-generating activities created by the *Carrera*. They became owners, consumers, and merchants of Asian commodities as they converted into brokers of global products, which had an exponentially expanding consumer base with their volume and variety. They typically consisted of affordable Asian goods of low to medium quality. Instead of being luxurious items or status symbols, the majority of these Asian imports were intended for general consumption and were tailored to meet the needs of a broad demographic in New Spain, Peru, and throughout Spanish America.

Acknowledgments. The author acknowledges the Office of the Chancellor of the University of the Philippines Diliman, through the Office of the Vice Chancellor for Research and Development, for funding support through the PhD Incentive Award.

References

Alonso, L. [2009] *El costo del imperio asiático: la formación colonial de las Islas Filipinas bajo dominio español, 1565-1800*. México: Instituto Mora, Universidade da Coruña.

Alonso, L. [2013] "E la nave va: economía, fiscalidad e inflación en las regulaciones de la carrera de la Mar del Sur, 1565–1604", in *Un océano de seda y plata: el universo económico del Galeón de Manila*. Seville: Consejo Superior de Investigaciones Científicas:25-84.

Alonso, L. [2019] "El 'alivio de las Indias:' la real hacienda Filipina, 1565–1800", *Estudis: Revista de Historia Moderna* 45:35-67.

Álvarez de Abreu, A. and C. Yuste [1977] *Extracto historial del comercio entre China, Filipinas y Nueva España (1736)*. México: Instituto Mexicano de Comercio Exterior.

Álvarez F. [1993] *Galeón de Acapulco: el viaje de la misericordia de Díos, relato histórico (1620)*. Madrid: Ediciones Polifemo.

Archivo General de Indias (AGI) [1591-1603] "Caja de Acapulco, cuentas de Real Hacienda [Treasury of Acapulco, Accounts of the Royal Treasury]", *Contaduría*:898-901.

Archivo General de Indias (AGI) [1604-1632] "Caja de Acapulco, cuentas de Real Hacienda [Treasury of Acapulco, Accounts of the Royal Treasury]", *Contaduría*:902-904.

Bonialian, M. [2012] "El pacífico hispanoamericano: política y comercio asiático en el imperio español (1680-1784)", PhD dissertation. El Colegio de Mexico AC.

Bonialian, M. [2014] "La contratación de la China por América colonial a principios del siglo XVII: la mirada de Francisco Valverde de Mercado, Gobernador de Panamá", *Boletín del Instituto de Historia Argentina y Americana Dr. Emilio Ravignani* 40:11-41.

Bonialian, M. [2021]. *La américa española: entre el pacífico y el atlántico; globalización mercantil y economía política, 1580-1840*. Mexico: El Colegio de Mexico AC.

Borah, W. [1975] *Comercio y navegación entre México y Perú en el siglo XVI*. Mexico: Instituto Mexicano de Comercio Exterior.

Borao, J. [2005] "La colonia de japoneses en Manila en el marco de las relaciones de Filipinas y Japón en los siglos XVI y XVII", *Cuadernos CANELA: revista anual de literatura, pensamiento e historia, metodología de la enseñanza del español como lengua extranjera y lingüística de la Confederación Académica Nipona, Española y Latinoamericana* 17:25-53.

Boyd-Bowman, P. [1973] "Spanish and European Textiles in sixteenth century Mexico", *The Americas* 29(3):334-358.

Castro, S. [2018] *Textiles in the Philippine colonial landscape: a lexicon and historical survey*. Quezon City: Ateneo de Manila University Press.

del Valle Pavón, G. [2005] "Los mercaderes de México y la transgresión de los límites al comercio pacífico en Nueva España, 1550–1620", *Revista de Historia Economica (Journal of Iberian and Latin American Economic History)* 23(S1):213-240.

Diaz-Trechuelo, M. [1964] "Philippine economic development plans, 1746-1779", *Philippine Studies* 12(2):203-231.

Flynn, D. and A. Giráldez [1995] "Born with a 'silver spoon': the origin of world trade in 1571", *Journal of World History* 6(2):201-221.

Flynn, D. and A. Giráldez [2008] "Born again: globalization's sixteenth century origins (Asian/global versus European dynamics)", *Pacific Economic Review* 13(3):359-387.

Gerritsen, A. and G. Riello [2015] "The global lives of things: material culture in the first global age", in *The global lives of things: the material culture of connections in the early modern world*. London and New York: Routledge:1-28.

Goode, C. [2017] "Merchant-bureaucrats, unwritten contracts, and fraud in the Manila Galleon Trade" in *Corruption in the Iberian empires: greed, custom, and colonial networks*. Albuquerque: University of New Mexico Press:171-196.

Hidalgo, P. [1995] *Encomienda, tributo y trabajo en Filipinas, 1570-1608*. Universidad Autónoma de Madrid.

Iaccarino, U. [2011] "The 'Galleon System' and Chinese trade in Manila at the turn of the 16th century", *Ming Qing Yanjiu* 16(01):95-128.

Iwasaki, F. [1992] *Extremo oriente y Perú en el siglos XVI*. Madrid: Mapfre.

McCarthy, W. [1993]. "Between policy and prerogative: malfeasance in the inspection of the Manila galleons at Acapulco, 1637", *Colonial Latin American Historical Review* 2(2):163-183.

Obispado, K. [2021] "The Pacific sailors: global workers at and on the edge of the Spanish empire (1580s-1640s)", PhD dissertation. El Colegio de México.

Obispado, K. [2023a] "The case of the dead sailors and the things they left: a microhistorical analysis of the Carrera del Pacífico in the sixteenth century", *Entre Europa y América: el mar y la primera globalización, Historia Medieval y Moderna* 96:199-212.

Obispado, K. [2023b] "Textiles and other trade goods: the Philippines in the sixteenth-century global trade", in *Plural Entanglements: Philippine Studie*s. Quezon City: Ateneo de Manila University Press:58-83.

Oropeza, D. [2007] "Los 'indios chinos' en la Nueva España: la inmigración de la nao de China, 1565–1700", PhD dissertation. El Colegio de México.

Ortigosa, J. and Y. Lee [2018] "Comercio y cargazón en el Galeón de Manila: el Santo Cristo", *Boletín americanista* 76:115-133.

Pinto, P. [2014] "Manila Macao and Chinese networks in south China Sea", *Anais de História de Além-Mar* 15:79-100.

Qing, L. [2018] "Trade and its historical trend between China and the Philippines in the late Ming Dynasty: an analysis on the almojarifazgo data", *Researches in Chinese Economic History* 3:171-181.

Recopilación de leyes, de los reynos de las Indias [1973]. *De las Indias mandadas imprimir y publicar por la Magestad Católica del Rey don Carlos II Nuestro Señor*. España: Madrid Cultura Hispánica.

Schurz, W. [1918] "Mexico, Peru, and the Manila galleon", *Hispanic American Historical Review* 1(4):389-402.

Schurz, W. [1939] *The Manila galleon*. EP Dutton.

Seijas, T. [2014]. Asian slaves in colonial Mexico: from Chinos to Indians. Cambridge University Press.

Tremml, B. [2012] "When political economies meet: Spain, China and Japan in Manila, ca. 1571-1644", PhD dissertation. University of Vienna.

Trusted, M. [2013] "Survivors of a shipwreck: ivories from a Manila Galleon of 1601", *Hispanic Research Journal* 14(5):446-462.

Villiers, J. [1980] "Silk and silver: Macau, Manila and trade in the China seas in the sixteenth century (a lecture delivered to the Hong Kong branch of the Royal Asiatic Society at the Hong Kong Club)", *Journal of the Hong Kong Branch of the Royal Asiatic Society* 20:66-80.

Yuste, C. [1984] *El comercio de la Nueva España con Filipinas, 1590-1785*. Instituto Nacional de Antropología e Historia, Departamento de Investigaciones Históricas.

Zapatero, A. [2012] "Un ejemplo de mundialización: el movimiento de biombos desde el Pacífico hasta el Atlántico (s. XVII-XVIII)", *Anuario de Estudios Americanos* 69(1): 31-62.

APPENDIX. Manila-Acapulco Trade Goods in 1592 based on the ten percent almojarifazgo charged to the crew

Spanish Terms	Item	Description	Unit and Quantity	Price per unit					
				IN ORO COMUN			IN ORO DE MINAS		
				Peso	Tomín	Grano	Peso	Tomín	Grano
A. Food									
azucar	sugar		4 arroba				5		
canela	cinnamon		16 cates				1		
B. Lead									
albayalde	white lead		1 arroba						
C. Personal Items									
abalorio	glass bead		5 cate	5					
aguja	needle								
abanico	fan		6 pieces						
abanillo comun	fan	common	360 dozens					3	
peine de tortuga	comb	made of tortoiseshell	250 dozens					2	
peine pintado de palo	comb	painted, made of wood	200 dozen					2	
peine de palo	comb	made of wood	300 dozens					2	
rosario	rosary		50 pieces					4	
D. Home Essentials									
campanilla pequeña	bell	small	2 pieces						
lanterna pequeña	lantern	small	2 pieces					4	
cajita dorada	box	gilded	14 pieces					4	
cajita pequeña comun	box	small, common	2 pieces						
cajita pintada	box	little, painted	8 pieces		2				
cajuela de aguja	needle case		24 pieces					2	

APPENDIX. Manila-Acapulco Trade Goods in 1592 based on the ten percent almojarifazgo charged to the crew (continued)

Spanish Terms	Item	Description	Unit and Quantity	Price per unit					
				IN ORO COMUN			IN ORO DE MINAS		
				Peso	Tomín	Grano	Peso	Tomín	Grano
cajuela dorada chica	trunk	small, golden	5 pieces					4	
escribania dorada	writing desk	gilded	1 piece				2		
escripuertorio	escritoire		1 piece				4		
escripuertorio dorado	escritoire	gold	1 piece						
piezas de Canton	pieces	Cantonese	40 pieces				5		
bacia mediana de metal	basin	medium, made of metal	11 pieces				2		
bacinilla	basin		7.5 dozens				1	6	
tachos y bacinilla de metal	pail and chamber pot	made of metal	85 pieces					4	
bandeja de palo	tray	made of wood	6 pieces					4	
bandeja de palo	tray	made of wood	6 pieces					2	
rodella	buckler		6 pieces					2	
escudilla	bowl								
escudilla basta	bowl	coarse							
escudilla chiquita de bejuco	bowl	small, made of bejuco	2 dozen				1	2	
escudilla ordinaria	bowl	ordinary	83 dozens					6	
cuchara de concha	spoon	similar in shape to seashell	610 dozens					2	
cuchara de nacar	spoon	decorated with mother of pearl	200 dozens					2	
limeta	long neck bottle		4 dozens and 2 pieces				1	7	
limeta chica	long neck bottle	small	110 dozens				1	6	

APPENDIX. Manila-Acapulco Trade Goods in 1592 based on the ten percent almojarifazgo charged to the crew (continued)

Spanish Terms	Item	Description	Unit and Quantity	Price per unit					
				IN ORO COMUN			IN ORO DE MINAS		
				Peso	Tomín	Grano	Peso	Tomín	Grano
loza basta	earthenware	coarse	900 dozens					6	
loza fina	china	fine	33 dozens and 4 pieces				2		
loza	earthenware		66 dozens				1	6	
loza	earthenware		68 bundles						
loza comun	earthenware	common	131.5 dozens					6	
loza comun, plato y escudilla	earthenware, plate, and bowl	common	89 dozens and 4 pieces	1	2			6	
loza, plato, etc	crockery, plate, etc.		354 dozens					1	
loza de Macan	earthenware	Macanese	6 dozens				2		
loza de Macan	earthenware	Macanese							
loza dorada	china	gilded	33 dozens				1	2	
almires	mortar		2 pieces					6	
pires	saucer		150 dozens					2	
plato	plate		16 dozens and 8 pieces					6	
plato chico basto	plate	small, coarse	1 dozen and 8 pieces					6	
plato chico comun	plate	small, common	100 dozens					6	
plato de baccina comun	basin plate	common	34 dozens and 10 pieces				1	6	
plato de baccina basta	basin plate	coarse	18 dozens and 4 pieces					1	
plato de baccina basta	basin plate	coarse	100 dozens				1	6	

APPENDIX. Manila-Acapulco Trade Goods in 1592 based on the ten percent almojarifazgo charged to the crew (continued)

Spanish Terms	Item	Description	Unit and Quantity	Price per unit					
				IN ORO COMUN			IN ORO DE MINAS		
				Peso	Tomín	Grano	Peso	Tomín	Grano
plato de baccina de losa basta	earthenware basin plate	coarse	125 dozens				1	6	
plato de loza comun	earthenware plate	common	96 dozens					10	
plato y escudilla fino de Macan	plate and bowl	Macanese, fine	25 dozens	2					
plato y escudilla basta	plate and bowl	coarse	83.5 dozens					6	
caracol	conch shells		126 dozens				1		
E. Animal									
pajara	female bird		2 pieces				3	4	
F. Slave									
esclavo	slave	Simon, black	1 person				150		
esclavo	slave	Hernando, age 18	1 person				120		
G. Raw Materials									
hilo	thread		97 cates					4	
hilo	thread		14 cates					4	
hilo al reves	thread	reverse	3 cates						
hilo bilado de oro y plata	thread	twisted, with gold and silver	2 cates	6					
hilo blanco	thread	white	50 cates	1	4			4	
hilo blanco	thread	white	18 cates	1	4				
algodón	cotton thread		10 libras				1		
algodón	cotton		1.75 quintals				15		
algodón	cotton		1 arroba				3		

APPENDIX. Manila-Acapulco Trade Goods in 1592 based on the ten percent almojarifazgo charged to the crew (continued)

Spanish Terms	Item	Description	Unit and Quantity	Price per unit					
				IN ORO COMUN			IN ORO DE MINAS		
				Peso	Tomín	Grano	Peso	Tomín	Grano
hilo de algodón	cotton thread		1 cate		6				
hilo gordo blanco	thread	thick, white	20 cates					4	
flusco de hilo	yarn	loose	1 libra	1					
flusco de seda blanca	silk	white, skein	105 libras: 3 dozen/1p	1					
seda cruda	silk	raw	12 cates				1	3	
seda floja	silk	floss	9.5 cates				1		
seda floja comun	silk	ordinary, floss	32.5 cates				1		
seda floja de color	silk	colored, floss	47.5 cates	2	4		1		
seda negra torcida	silk	black, twisted	5 cates				1	3	
seda torcida	silk	twisted	24 cates				1	3	
seda torcida blanca	silk	white, twisted	cate: 1/2 pico				1	3	
seda torcida de oro	silk	gold, twisted	16.5 cates	1	3				
seda torcida de color al reves	silk	twisted	cate: 1/2 pico				1	3	
H. Fashion Accessory									
liga comun	garter	ordinary	30 pairs					2	
liga de Japon	garter	Japanese	5 pieces					4	
liga sin punta	garter	tipless	4 piesce					2	
reata	string		14 pieces				1		
pasamano de plata	braid	with silver thread	2 cates	10					
pasamano de seda y oro	braid	with silk and gold thread	5 cates	10					
pasamano negro de hilo	tassle	with black thread	1 cate					4	

APPENDIX. Manila-Acapulco Trade Goods in 1592 based on the ten percent almojarifazgo charged to the crew (continued)

Spanish Terms	Item	Description	Unit and Quantity	Price per unit					
				IN ORO COMUN			IN ORO DE MINAS		
				Peso	Tomín	Grano	Peso	Tomín	Grano
I. Woven fabric									
medriñaque	abaca cloth		50 pieces					3	
medriñaque	abaca cloth		100 pieces					4	
sinabafa	cotton fabric	Philippine	7 pieces		9				
sinabafa de chincheo	cotton fabric	from Chincheo	333 pieces						
sinabafa negra	cotton fabric	black	1 piece						
caniquí angosto	muslin	fine	16 pieces					2	
damasco	damask		67 pieces				5		
damasco	damask	11 varas	47 pieces				5		
damasco	damask		2.5 dozens				6		
damasco de color de arretes	damask	color of arretes	2 pieces	22					
damasco de color labor de China	damask	Chinese, colored	22 pieces	9					
damasco manchado comun	damask	common, spotted	3 pieces						
damasquillo comun	damask-like cloth	ordinary, damask-like	19 pieces	3			1	4	
manta	cotton cloth, fabric, or blanket	8 varas							
manta ancha de ojo de perdiz	cotton cloth, fabric, or blanket	bird's eye patterned	97 pieces	1	4			6	
manta angosta de sangley	cotton cloth, fabric, or blanket	Sangley, thin	114 pieces		2			2	
manta angosta de sangley	cotton cloth, fabric, or blanket	Sangley, thin	50 pieces					1	

APPENDIX. Manila-Acapulco Trade Goods in 1592 based on the ten percent almojarifazgo charged to the crew (continued)

Spanish Terms	Item	Description	Unit and Quantity	Price per unit					
				IN ORO COMUN			IN ORO DE MINAS		
				Peso	Tomín	Grano	Peso	Tomín	Grano
manta azul	cotton cloth, fabric, or blanket	blue	25 pieces					2	
manta blanca y negra de Canton	cotton cloth, fabric, or blanket	Cantonese, white and black	7 pieces					5	
manta blanca	cotton cloth, fabric, or blanket	white	10 pieces					5	
manta blanca	cotton cloth, fabric, or blanket	white	39 boxes					3	
manta cruda	cotton cloth, fabric, or blanket	crude	15 pieces	1	4			5	
manta cruda	cotton cloth, fabric, or blanket	crude	37 pieces					6	
manta cruda angosta de chincheo	cotton cloth, fabric, or blanket	from Chincheo, crude, lightweight	21 pieces					6	
manta cruda de Canton	cotton cloth, fabric, or blanket	crude	613 pieces					5	
manta cruda de chincheo	cotton cloth, fabric, or blanket	from Chincheo, crude	10 pieces						
manta de algodón	cotton cloth, fabric, or blanket		35 pieces						
manta de Canton	cotton cloth, fabric, or blanket	Cantonese	3087 pieces		9			5	
manta de Canton	cotton cloth, fabric, or blanket	Cantonese, 8 libras	92 pieces					5	
manta de Canton	cotton cloth, fabric, or blanket	Cantonese, 8 varas	400 pieces					5	
manta de chincheo	cotton cloth, fabric, or blanket	from Chincheo	3578 pieces		9			6	

APPENDIX. Manila-Acapulco Trade Goods in 1592 based on the ten percent almojarifazgo charged to the crew (continued)

Spanish Terms	Item	Description	Unit and Quantity	Price per unit					
				IN ORO COMUN			IN ORO DE MINAS		
				Peso	Tomín	Grano	Peso	Tomín	Grano
manta de chincheo	cotton cloth, fabric, or blanket	from Chincheo, 11 to 12 varas	1500 pieces		9				
manta de chincheo comun	cotton cloth, fabric, or blanket	from Chincheo, common	100 pieces						
manta de chincheo que son angostas de Sanglei	cotton cloth, fabric, or blanket	from Chincheo, thin	170 pieces					2	
manta de Chincheo y Canton de las ordinarias	cotton cloth, fabric, or blanket	from Chincheo and Canton, ordinary	140 pieces					5	
manta listada	cotton cloth, fabric, or blanket	striped	100 pieces					2	
manta listada de color	cotton cloth, fabric, or blanket	color striped	2 pieces					2	
manta de cordoncillo	cotton cloth, fabric, or blanket	laced	240 pieces		9			6	
manta de hierbesilla	cotton cloth, fabric, or blanket		60 pieces						
manta de Moro	cotton cloth, fabric, or blanket	Moro	88 pieces		4			2	
manta de Sangley	cotton cloth, fabric, or blanket	Sangley	30 pieces		9				
manta de Sangley	cotton cloth, fabric, or blanket	Sangley	3 pieces	1					
manta de seda	silk cloth		4 pieces and more				2		
manta fina de Lanquin	cotton cloth, fabric, or blanket	from Lanquin, fine	6 pieces				1		
manta morada	cotton cloth, fabric, or blanket	maroon	2 pieces					2	

APPENDIX. Manila-Acapulco Trade Goods in 1592 based on the ten percent almojarifazgo charged to the crew (continued)

Spanish Terms	Item	Description	Unit and Quantity	Price per unit					
				IN ORO COMUN			IN ORO DE MINAS		
				Peso	Tomín	Grano	Peso	Tomín	Grano
manta negro de cordoncillo	cotton cloth, fabric, or blanket	black, corded	17 pieces						
mantas listada de azul y blanco	cotton cloth, fabric, or blanket	blue and white striped	17 pieces						
manta pintada	cotton cloth, fabric, or blanket	painted	10 pieces		4				
raso amarillo	satin	yellow	1 piece						
raso blanco	satin	white	4 pieces	9					
raso de Canton	satin	Cantonese	20 pieces				5		
raso de color	satin	colored	50 pieces	9					
raso de color arrollado	satin		3 pieces				5		
raso de color fino de Lanquin	satin	Lanquin, fine-colored	68 pieces				6		
raso negro de Canton	satin	Cantonese, black	9 pieces	11			5		
raso negro de Lanquin	satin	from Lanquin, black	6 pieces	13					
raso negro y de color de Canton	satin	Cantonese, black, colored	30 pieces	11			5		
seda blanca	silk	white	10 cates	3					
seda de color	silk	colored	60 cates				1		
seda de Japon	silk	Japanese	7 pieces				2		
seda de Japon sencillo	silk	ordinary	6 pieces					2	
brocadillo falso comun encarnada	brocade	common imitation, flesh colored	2 pieces	5					
burato	wool or silk fabric		1 piece				3		

APPENDIX. Manila-Acapulco Trade Goods in 1592 based on the ten percent almojarifazgo charged to the crew (continued)

Spanish Terms	Item	Description	Unit and Quantity	Price per unit					
				IN ORO COMUN			IN ORO DE MINAS		
				Peso	Tomín	Grano	Peso	Tomín	Grano
telilla comun	woolen cloth	light, ordinary	6 pieces	5			1	4	
gorgorán	grogram		93 pieces	5			1	6	
gorgorán ajedresado comun	grogram	common, checkered	1 piece	5					
gorgorán de color	grogram	colored	40 pieces				1	6	
gorgorán negro	grogram	black	6 pieces				2	4	
gorgorán negro comun	grogram	ordinary, black	2 pieces						
gorgorán tornasol	grogram	ordinary, iridescent	1 piece						
tafetan comun	taffeta	ordinary	229 pieces	3			1	6	
tafetan blanco comun	taffeta	white, ordinary	2 chasubles	8					
tafetan blanco comun	taffeta	8 varas, white, ordinary	4 pieces	3					
tafetan de color	taffeta	colored	30 pieces					6	
tafetan de color	taffeta	colored	60 pieces				1	6	
tafetan de color común	taffeta	7 and 1/2 varas, colored, ordinary	6 pieces						
J. Head Covering									
mantilla listada de color	veil or shawl	color-striped							
mantilla listada	veil or shawl	striped	6 pieces					2	
mantilla pintada	veil or shawl	painted	69 pieces					2	
toca	headcloth	7 varas y media	40 pieces		5				
toca	headcloth	7 varas	10 pieces		5				
toca	headcloth	6 varas	5 pieces						
toca ordinaria	headcloth	ordinary	1032 pieces					2	6

APPENDIX. Manila-Acapulco Trade Goods in 1592 based on the ten percent almojarifazgo charged to the crew (continued)

Spanish Terms	Item	Description	Unit and Quantity	Price per unit					
				IN ORO COMUN			IN ORO DE MINAS		
				Peso	Tomín	Grano	Peso	Tomín	Grano
toca	headcloth		700 pieces					2	
toca cruda	headcloth	unrefined	116 pieces					2	6
toca cruda	headcloth	unrefined	20 pieces					2	
toca cruda ordinaria	headcloth	unrefined, ordinary	119 pieces					2	6
toca de espumilla	headcloth	loosely woven	2 pieces					2	6
toca negra ordinaria	headcloth	black, ordinary	3 pieces					2	6
toca de red	hairnet		126 pieces		4			3	
toca de red	hairnet		20 pieces					6	
toquilla blanca	headscarf	small, white	2 pieces					2	6
K. Personal Cloth									
pañito de nariz de manta	cotton handkerchief		17 pieces					2	
paño de mano	hand cloth		79 pieces					2	
paño de mano de manta	cotton hand cloth		10 pieces	1				2	
pañuelo de rostro	sweat cloth								
pañito/pañuelo de nariz	handkerchief		24 pieces					1	
L. Garments									
cuello llano	collar	plain	7 pieces					2	
manga de raso	sleeves	made of satin	9 pieces				1	2	
camisa de manta	blouse		23 pieces					7	
camisa llana	shirt	plain	21 pieces					7	
camisilla	undershirt		5 pieces		2				

APPENDIX. Manila-Acapulco Trade Goods in 1592 based on the ten percent almojarifazgo charged to the crew (continued)

Spanish Terms	Item	Description	Unit and Quantity	Price per unit					
				IN ORO COMUN			IN ORO DE MINAS		
				Peso	Tomín	Grano	Peso	Tomín	Grano
ropilla de raso negro	double-sleeved short jacket	made of black satin	3 pieces				1		
ropilla de lienzo sin mangas	sleeveless vestment	made of linen	5 pieces						
cangan	jacket or short-sleeved coat		40 pieces					4	
cangan azul	jacket or short-sleeved coat	blue	9 pieces					4	
jubon de manta	waistcoat	made of cotton cloth	1 piece				1		
jubon de tela	jerkin	made of linen	3 pieces				1	2	
jubon de lienzo sin manga	sleeveless jerkin	made of linen					1		
frontal con su frontalera	frontal	with ornaments							
saya colchada	skirt	quilted	5 pieces				4		
saya colchada de sinabafa	skirt	quilted, made of sinabafa	16 pieces						
calzon	breeches		20 pieces					3	
calzon blanco	breeches	white	4 pieces					3	
calzon de raso negro	breeches	black	2 pieces				3		
M. Footwear									
medias de algodón	stockings	cotton	98 pairs					2	
medias de lienzo	stockings	linen	5 pieces					1	
medias de seda	stockings	silk	6 pairs				2		
medias de seda negra	stockings	black, silk	3 pairs						
zapatos y pantufos	shoes and slippers		1 dozen						

APPENDIX. Manila-Acapulco Trade Goods in 1592 based on the ten percent almojarifazgo charged to the crew (continued)

Spanish Terms	Item	Description	Unit and Quantity	Price per unit					
				IN ORO COMUN			IN ORO DE MINAS		
				Peso	Tomín	Grano	Peso	Tomín	Grano
N. Bedding and Curtain									
colcha blanca	bedcover	white	1 piece				4		
cobija	blanket								
cobija blanca bordada	blanket	white, embroidered	2 pieces				7		
sabana	bedsheet		4 pieces					2	
sabana de manta	bedsheet	made of cotton cloth	2 pieces					4	
sobrecama	beadspread		3 pieces				6		
sobrecama	beadspread		6 pieces				15		
sobrecama colchada	beadspread	quilted	2 pieces						
sobrecama de Damasco	beadspread	damask							
sobrecama de Damasquillo	beadspread	made of damask-like cloth	3 pieces				4		
sobrecama de manta	beadspread	made of manta	1 piece				2		
almohada de manta	pillow cover	made of manta	4 pieces					2	
pabellon blanco	bed canopy	white	1 piece				3		
pabellon de hierbesilla	bed canopy		2 pieces						
pabellon de manta	bed canopy	made of manta	1 piece				3		
pabellon de tafetan	bed canopy	made of taffeta	2 pieces				15		
O. Others									
varias mercadurias	various merchandise								
otros sin registros	unregistered								

Source: Archivo General de Indias (AGI) [1592] Caja de Acapulco, Cuentas de Real Hacienda [Treasury of Acapulco, Accounts of the Royal Treasury]. Contaduría 898.
Note:
1 oro comun = 9 reales or 306 maravedis
1 oro de mina = 13 and 1/4 reales or 450 maravedis
1 real = 34 maravedis

The Philippine Review of Economics

ERRATA

The Philippine Review of Economics [2023] 60(1):10-18. DOI:10.37907/2ERP3202J

Errata to **Introduction to the symposium on care economy**

Maria S. Floro
American University

Elizabeth M. King
The Brookings Institution

On page 15, last line "...[National Statistics Authority 2023]." should read as "… [Philippine Statistics Authority 2023].

Page 18 in the reference list

National Statistics Authority [2023] "National migration survey", https://psa.gov.ph/tags/national-migration-survey. Accessed June 2023."

should read as

"Philippine Statistics Authority [2023] "National migration survey", https://psa.gov.ph/tags/national-migration-survey. Accessed June 2023.

Published online June 30, 2023.

The Philippine Economic Society

Founded 1961

The Philippine Economic Society (PES) was established in August 1962 as a nonstock, nonprofit professional organization of economists.

Over the years, the PES has served as one of the strongest networks of economists in the academe, government, and business sector.

Recognized in the international community of professional economic associations and a founding member of the Federation of ASEAN Economic Associations (FAEA), the PES continuously provides a venue for open and free discussions of a wide range of policy issues through its conferences and symposia.

Through its journal, the *Philippine Review of Economics* (PRE), which is jointly published with the UP School of Economics, the Society performs a major role in improving the standard of economic research in the country and in disseminating new research findings.

At present the society enjoys the membership of some 800 economists and professionals from the academe, government, and private sector.

- Lifetime Membership – Any regular member who pays the lifetime membership dues shall be granted lifetime membership and shall have the rights, privileges, and responsibilities of a regular member, except for the payment of the annual dues.
- Regular Membership – Limited to individuals 21 years of age or older who have obtained at least a bachelor's degree in economics, or who, in the opinion of the Board of Directors, have shown sufficient familiarity and understanding of the science of economics to warrant admission to the Society. Candidates who have been accepted shall become members of the Society only upon payment of annual dues for the current year.
- Junior Membership – This is reserved for full-time college or graduate students majoring in economics. Affiliation for junior membership is coursed through the Junior Philippine Economic Society (JPES).

For more information, visit: www.phileconsociety.org.