Theoretical and Applied Economics
Volume XXV (2018), No. 2(615), Summer, pp. 69-104

How to build an economy free of recession and stagnation: results from a multi-commodity macro model

Samuel MENG
University of New England, Australia
xmeng4@une.edu.au

Abstract. *Aggregate macro models used in existing studies have insufficient details to shed any light on the market saturation phenomenon and thus have failed to uncover the fundamental cause of economic recessions. A multi-commodity macroeconomic model is constructed reflecting three simple but revolutionary axioms, which summarize the important but often overlooked facts in an economy. The results from the model shows that, without product innovation, the economy will reach a consumption/income ceiling so that recessions are inevitable. In order to build an economy free of recession, the paper calls for stimulating innovation through a thorough revision of the current patent laws.*

Keywords: economic recessions, multi-commodity macro model, product innovation, patent laws, economic growth.

JEL Classification: E10, E20, E32.

1. Introduction

Economic recessions continue to haunt mankind and the effects of large economic recessions can be quite devastating. During the Great Depression of the 1930s, the output of several Western nations fell by 25-30% in the period between 1929 and 1932-33. In the United States (US), the unemployment rate peaked at 25% in 1932-33, and it took almost a decade for US output to return to its pre-1929 level.

More recently, the Global Financial Crisis (GFC) and the ongoing European Debt Crisis are striking examples of the recurrence of economic recessions. During the GFC, the US Dow Jones Industrial Average index dropped, from its peak of more than 14,000 points in October 2007, to a trough of less than 6,600 points in March 2009. In the last quarter of 2008, US quarterly real GDP decreased by 8.9%. The unemployment rate in the US increased from 4.6% in 2007 to 10.1% in 2010. As a result of the European Debt Crisis, many nations in Europe also experienced high levels of unemployment. According to the European Commission (2013), unemployment rates in 2012 were 25.0% in Spain, 24.3% in Greece, 15.9% in Portugal, and 14.9% in Cyprus.

The negative impact of economic recession has triggered a considerable amount of research. However, the studies so far have not pinpointed the causes of recessions and provide no effective method to avoid a recession. This partly may be due to the loose definition of economic recession which fails to reveal the key feature of a recession. In order to discuss the way to avoid economic recession or stagnation, it is necessary to clarify the economic recession or stagnation referred in this paper.

The term 'economic recession' or 'economic stagnation' is often used loosely to describe a period of slowing down or stagnancy of economic activities. It is often reported that National Bureau of Economic Research (NBER) defines a recession as decreases in GDP in two or more consecutive quarters, but NBER website defines a recession as 'a significant decline in economic activity spread across the economy, lasting more than a few months, normally visible in real GDP, real income, employment, industrial production, and wholesale-retail sales'. These definitions do not give a clear distinction between economic recession and economic fluctuation, so a recession can generally be interpreted as a large and long economic fluctuation from the average growth trend.

In this paper, the author defines economic recession or stagnation according to the key feature and possibly the cause of an economic downturn. The economic downturns caused by supply-side factors such as wars and pandemics are normally associated with a shortage of commodity supply. This type of downturns are easy to understand and easy to solve: since the downturn is caused by limitation in resources, the recession can be ended by increasing resources for production. This type of fluctuations generally does not cause boom-bust cycle (e.g. a cyclone will cause a large decrease in economic activity. We may expect the economy will recover later but we generally do not expect the cyclone disaster will lead to a new boom in the economy). However, the recession in the absence of a supply shock is generally associated with stagnation of sales, large amount of unemployment, and unutilized capital. This type of recessions is hard to explain and so far no one has found a solution. It is the recessions with abundance of commodities that we are trying to address in this paper.

One may argue: how can your definition accommodate an economic recession like GFC? On the surface, the GFC resulted from the housing mortgage debt crisis so most people believe that the cause of the GFC is the problems in the financial system. Some blame Alan Greenspan for the relaxed financial regulation. However, to find out the underlying cause for the GFC, one needs to ask the following questions. Why were so many investors including banks competing fiercely to provide housing mortgage lending to such a degree that the lending requirements deteriorated to almost nothing? Given that the poor lending standard is an open secret long before 2008, why did the investors not sense the danger of such a low lending requirement early? The lack of regulation and the principal-agent problem of the financial intermediates are of course to blame, but the primary reason is that there is no other profitable investment opportunities – if there are other sunny spots, investors will not crowd into mortgage lending. The reality was, facing stagnation of sales prior to the GFC, investing in commodity production was a path to bankruptcy. With no opportunity of investing in production, the idle capital has to go to assets markets and housing mortgage. The lack of investment opportunity due to stagnation of sales also explains why the recovery of GFC took so long. Consequently, we can conclude that the GFC is fundamentally a recession due to stagnation of sales.

The relationship between economic recession (or business cycle) and economic growth is also worth further clarification. The common wisdom in macroeconomic circle is that an economy grows with a trend in the long run while the fluctuations around the trend in the short run form business cycles. As a result, studies on economic growth are generally about the long-run growth trend while studies on business cycles are about economic fluctuations in the short run. The author rejects this traditional disconnected view about economic growth and business cycle. General economic fluctuations (e.g. seasonal fluctuations, or even large fluctuations caused by extreme weather, wars, natural disasters) are generally short and have little impact on economic growth trend. However, economic recessions associated with abundance of commodities tend to be persistent for a long period and are likely to cause boom-bust cycles. These business cycles have significant impacts on economic growth trend in the long run. Who can deny the impact of long economic recessions like the Great Depression and the GFC on economic growth?

By isolating economic growth from business cycles, many economists also delink the supply and demand sides. Generally, economic growth is regarded as being caused by supply factors such as labour, capital, natural resources, and technological progress; on the other hand, business cycles are caused by demand side factors such as demand deficiency and lack of animal spirit proposed by Keynes. This kind of thinking in isolation fails to uncover the links between the factors on supply side and those on demand side.

This paper intends to link the supply side with demand side to solve both the problem of economic recession and the problems in economic growth. To uncover the fundamental cause of economic recessions, the author has proposed three axioms based on common wisdom, and has developed a multi-commodity macroeconomic model. The model preserves the Classical assumption for an economy – the perfectly competitive market – and can provide a simple and universal explanation for recurrent economic recession and for cyclical economic growth.

The remainder of the paper is organized as follows. Existing theories on economic recessions and growth are briefly reviewed and discussed in Section 2. Based on observations and reasoning in the real world, Section 3 identifies three axioms supported by real world observations. In Section 4, a static multi-commodity model is developed and, based on this model, a recursive dynamic model and an intertemporal model are also developed. Section 5 analysed the implications of modelling results and reveals the unique role of product innovation in both business cycles and economic growth. The reasons for product innovation scarcity and the ways of encouraging innovation are also discussed briefly in this section. Section 6 summarizes the paper and provides brief concluding comments.

2. Major existing theories on economic recessions and growth

There are many studies on economic recessions, so it is neither possible nor necessary to review them all in this paper. Instead, the author will review key studies on the basis of various schools of economic thought and focus only on the key concepts of these studies.

Classical economists (Old, New, or Neo) have great faith in the efficiency of market mechanisms and in perfectly competitive markets. They regard economic recessions as large natural economic fluctuations (e.g. Lucas, 1975; Kydland and Prescott, 1982; Plosser, 1989; Prescott, 1986). They believe that, if market forces were allowed to operate alone, economic recessions would be temporary or relatively short-lived. Consequently, they argue that government intervention is unnecessary.

Regarding economic growth, classical economic theory has evolved from Solow's exogenous growth model to Romer's endogenous growth model. In a Solow-style growth model (e.g. the AK model), the technology level is viewed as constant or exogenous, so the economic growth rate is determined by capital growth rate or saving rate. A Romer-style growth model contributes technological growth to the growth of human capital, or labour devoted to R&D, or the variety of products. Jones (1995) provides a detail argument that these endogenous growth models are not consistent with empirical data (e.g. TFP growth rate does not correlate to the growth of capital and/or labour, even not correlate to the share of labour devoted to R&D) and proposes a semi-endogenous growth model (economic growth depends on the parameter indicating how sensitive technological growth responds to the amount of labour devoted to R & D). More importantly, these models are unable to explain the cyclical economic growth pattern. Although the Schumpeterian-style endogenous growth model (e.g. Aghion and Howitt, 1992, 2006) can produce cyclical equilibria under some circumstances, these results are critically hinged on creative destruction or vertical innovations – innovations create better quality of productions of a similar product, e.g. a more powerful (higher speed) computers.

Classical economists tend to deny or ignore the important features of economic recession highlighted by Keynesian economists. The real business cycle model, for example, view economic recessions as economic fluctuations caused by supply side shocks like a decrease in productivity, a hike in oil prices, and a decrease in labour supply. This model explains none of the main features of an economic recession such as stagnant demand, unutilized capital, and high unemployment. These features indicate clearly that an economic recession is mainly a problem on the demand side: the economy has plenty of resources (e.g. unemployed capital and labour) and capacity to produce but the sales stagnate. Classical economists palely argued that the appearance of oversupply is due to over-production by high cost firms or to mismatch of production and consumption, and that high unemployment is due to inflexible wages or voluntary unemployment. Even if these far-fetched arguments are true and thus an economic recession can be viewed as an economic disequilibrium, the question to be answered is why an efficient market allows this disequilibrium to last for years or even for a decade? Classical economists have to admit that either the market is inefficient (so a disequilibrium can last for a decade) or a recession is not simply disequilibrium. In either way, classical economists will contradict their own belief.

By describing economic recessions as natural fluctuations, however, classical economists avoid the task of finding the causes of economic recessions. Instead, they focus on developing economic models and econometric estimations and choose to be indifferent to the economic and psychological damage of a recession on human beings. It is not surprising that the classical economics solution to economic recession – natural recovery – is unpopular with government and public alike.

Alternative explanation of economic recession is the underconsumptionist view. This view is expressed as early as in 1723. In his satire poem the Fable of the Bees, Mandeville (1723) argued that it was ‘the Vices’ such as intemperance, luxury and pride of man that led to the high consumption, which promoted manufacture and industry. So these vices are the base for prosperity and public happiness. Hutcheson (1750, p.66, p.61) admitted that a ‘small part of our consumption … is owing to our Vices’, but he thought an “equal consumption of manufactures, and encouragement of trade may [exist] without these Vices’. It is Adam Smith who totally defeated the argument of Mandeville. Smith (1776, p321) stated ‘That portion which he annually saves, as for the sake of profit it is immediately employed as a capital, is consumed in the same manner, and nearly in the same time too, but by a different set of people’.

However, the reasoning of Smith that savings are immediately invested and thus consumed overlooked the fact that investment is dependent on expected future consumption: without an increase in future consumption, the goods brought about by investment cannot be sold. This point was picked up by Malthus (1836), but Malthus’ concern was quickly dismissed by Ricardo (1951) and Mill (1844) because they thought people’s will to purchase was very seldom wanting where the power to purchase existed. Later, the underconsumptionist view was revived and developed by John Keyes and Karl Marx.

The apparent oversupply or underconsumption of commodities during an economic recession was explained by Keynes (1936) and further by his successors, labelled ‘Keynesian economists’ (either ‘old’, ‘orthodox’, ‘new’, or ‘post’). The main contribution of Keynesian economics to explaining economic recession was shown by the concept of deficiency of effective demand. Keynes attributed this deficiency to decreases in investment. He demonstrated that a decrease in investment would lead to a proportionally greater decrease in output through a multiplier effect. Keynes determined the most important causes of this investment shortage to be, first, a lack of ‘animal spirits’ (entrepreneurship), and second, the liquidity preference, or the speculative motive to hold cash in a world characterized by uncertainty (the ‘uncertainty argument’). On the factor market, Keynes attributed high unemployment during a recession to the fluctuations of expected profit (or ‘marginal efficiency of capital’ in Keynes’ words), resulting from unstable investment expenditure. The liquidity preference and uncertainty argument was further developed by post-Keynesian economists (e.g. Davidson, 1984, 1991), while microeconomic foundations for demand deficiency and unemployment were developed by new-Keynesian economists, including the ideas based on real and nominal wage rigidity, price rigidity, efficiency wages, etc. (e.g. Mankiw, 1985, 1989, Akerlof and Yellen, 1985; Romer, 1993).

Marx’s explanation of economic recession has been given little attention in economics literature, perhaps because of his radical idea of advocating class warfare. Nonetheless, there is an element of truth in the Marxist argument that warrants discussion here. Marxists determine that economic recession is caused by inequality. Their explanation is based on their observation of the behaviours of capitalists. In order to obtain as much as possible profit, capitalists tend to produce as much as possible and, on the other hand, try to push wages down and raise the rate of surplus value. As a result, the workers are unable to buy up the value they produced and this causes excess supply and inadequate aggregate demand.

Keynesian economists intuitively identified that the key features of an economic recession are depressed demand and high rates of unemployment. These features were, however, attributed to quite unusual factors, e.g. liquidity preference and the lack of entrepreneurship (by Post-Keynesian economists), and wage and price rigidity in an economy (by New-Keynesian economists). The question unanswered by Keynesian economists is why these factors exist during a recession but do not exist in an economic boom? In other words, Keynesian economists have not gone far enough to uncover the causes for liquidity preference and the lack of entrepreneurship. Keynesian economists discarded the long-standing assumption of Classical economics that perfectly competitive markets exist. By rejecting the existence of Adam Smith’s ‘invisible hand’ (i.e. the efficiency of the market), the solution by Keynesian economists became one of interventionism.

Marxists have highlighted the inequality issue. There is little doubt that income distribution inequality plays an important role in economic recessions and this role was fully demonstrated by the GFC. Large-scale lending by the rich to the poor boosted demand for housing and consumption but, when the rich sensed a risk of loan default and wanted their loans to be repaid, the resulting financial constraint on the poor, and thus the consequent decrease in final demand, pushed the economy into recession. While it is clear that income inequality is a contributing factor to economic recession, this inequality may not be a fundamental factor underpinning recession – otherwise the economy should stay in recession or stagnation because the unequal income distribution has not changed. However, history shows that every recession has moved to recovery and expansion stages. Moreover, in considering the large production capacity in the modern global economy, there is always a possibility for overproduction and thus deficiency of demand, even if income is equally distributed and everyone has sufficient income to buy what they want. From this point of view, income inequality can aggravate or accelerate a recession, but it is not the fundamental cause.

Although both Keynesian and Marxism provided some explanation to underconsumption, they failed to gain the orthodox status. This may be because its underconsumptionist's view is largely rejected by history: it is proven that consumption does keep going up over time and economic recessions are eventually over. However, one should not discard the whole underconsumptionist's theory without a careful reasoning. One can say that a person's desire for all kinds of goods (i.e. overall consumption) is unlimited and thus overall underconsumpion is implausible, but a person's desire for any type of goods may be limited and thus may cause underconsumption or overproduction of any type of goods. This is evident by the phenomenon of market saturation. The paradox between the limited consumption of each type of commodity and the unlimited overall consumption rests on the increasing number of goods: as time passes by, product innovation keeps bringing us new goods and services such as i-phones, driverless cars, and space travel. This also explains the paradox between the apparent underconsumption (i.e. stagnation of sales) during an economic recession and forever increasing consumption in the long run. These paradoxes suggests that it is important to study the impact of product innovation on our consumption and thus on economic recessions.

From the above discussion it can be said that previous studies of economic recession hold some elements of truth. However, they have failed to uncover the fundamental cause of recession, and thus have failed to provide a satisfactory solution. As long as economic recessions continue to occur, the search for answers must also continue.

3. Axioms to overcome the shortcomings in existing economic theories

The foregone discussion shows that the existing theories view economic growth and business cycles from different angles but all these theories have drawbacks. To overcome these, this section proposes three simple but revolutionary axioms, which summarize the fundamental but overlooked facts in an economy. While these axioms may look unconventional, they are comprehendible because they are based on real world observations and logical reasoning. The focus of this section is to discuss these axioms.

Axiom 1 (consumption satiety): every individual has a satiation point in consumption of any commodity. Consumption over this satiation point is harmful.

Most economists are against the idea of consumption satiety or consumption ceiling. It can be argued that the assumption of a consumption ceiling is inappropriate because people's desire for goods and services is unlimited, or simply that consumption increase over time without any ceiling. In accepted preference and utility theory, it is assumed that consumers always prefer greater quantities of any kind of commodity without limit. That is, the greater the quantity of commodities consumed, the higher the utility will be. This view held even for Keynes, who believed in demand deficiency but never thought of consumption deficiency, instead, he attributed the deficiency of effective demand to the deficiency in investment demand.

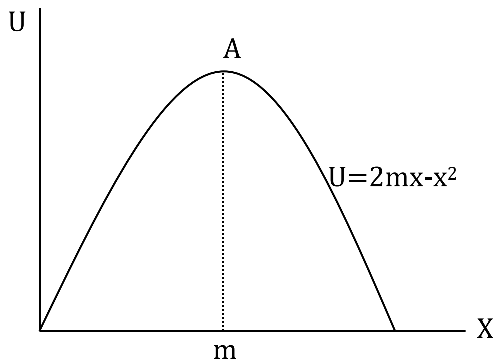
We should, however, not just think about consumption loosely. Next we focus on the case of one person's consumption of one commodity, for example, one's consumption (i.e. eating) of ice cream. Most people like ice cream and, generally speaking, the more ice cream consumed, the higher utility will be achieved. Eating too much ice cream will, however, lead to vomiting or stomach pains. Even for a harmless commodity like shoes and clothes, one must have limit on the quantity – given a large number of shoes and clothes, one need a lot of storage space and a lot of time to find the one you want to put on. For services like concert and massage, listening to music for too long in a day would damage your hearing (in this case music becomes noise), too much massage would damage your skin (due to friction). These examples can be generalized to any goods and services (the author challenges the reader to name a commodity otherwise), and suggests that overconsumption can be a burden for a consumer.

The above example demonstrates that a satiation point exists for the consumption of any commodity (here the concept of consumption is strictly defined as the use-up of goods or services)[(1)]. If the amount of consumption surpasses the satiation point, the utility from consumption will decline. To embody the satiation point in the utility function, a parabolic utility function can be proposed:

$$U(x) = 2mx - x^2$$

This utility function is illustrated in Figure 1.

Figure 1. *Utility of a representative individual with consumption*



When the amount of consumption of *x* is less than *m*, the utility achieved increases as the consumption increases. Once *x* is greater than *m*, however, further increases in consumption will result in lower utility. For a rational consumer, the maximum consumption of *x* is *m* and the maximum utility is achieved at point A.

Since each commodity we consume has a satiation point, adding up all satiation amounts of each commodity, we can have a consumption ceiling for any consumer or household. Adding up all the consumption ceilings of all households in an economy, we can have a consumption ceiling for the economy (we can consider the consumption ceiling per capita to simplify the case of changing population). In short, The one-commodity case may be generalized and applied to the real world of multiple types of commodities (there is no fallacy of composition here because the satiation point can be applied to each commodity and to each household). Since an individual needs different types of commodities (and services) to satisfy his or her needs, the following utility function can be used for a general case:

$$U = U(x_1, x_{2,} ... x_n) = \sum_{i=1}^{n} \alpha_i (2m_i x_i - x_i^{\,2})$$

Where α_i is the weight of consuming good x_i in the individual's utility.

The number of commodity type *n* in the above equation is very important because it is the key to understanding the seeming contradiction between consumption satiety and unlimited human desire. A human being cannot consume unlimited amount of any good, but his/her desire can be unlimited when the varieties of goods are abundant and/or when the varieties increase thanks to the new products being invented over time.

Consumption satiation is in fact not a new concept. The consumption ceiling caused by consumption satiation points has been studied by other researchers such as Fisk (1962), Stent and Webb (1975), Pasinetti (1981), Witt (2001), Anderson (2001), Aoki and Yoshikawa (2002), Saviotti and Pyka (2013) and Chai and Moneta (2014). In fact, Pasinetti proposed a satiation hypothesis similar to Axiom 1: 'there is no commodity for which any individual's consumption can be increased indefinitely. An upper saturation level exists for all types of goods and services although at different levels of real income' (Pasinetti, 1981, p77). He classified goods into three types. One is necessities for physiological needs (e.g. food) which have an Engel curve with expenditure growing slower and reaching a maximum point as income increases, the second is other normal goods which have an Engel curve with expenditure accelerated and then decelerated and reaching a maximum point as income increases; the third one is the inferior goods which have an Engel curve with expenditure increasing to the maximum point and then decreasing as income increases. The difference presented in Axiom 1 is that all goods are inferior goods at different level of real income. Moreover, although the implication of satiation phenomenon on economic growth as well as innovation activity is studied by many researchers (e.g. Pasinetti, 1981; Saviotti, 2001; Witt, 2001; Ruprecht, 2005), this paper will contribute more along this line by using Axiom 1 to study business cycle and thus shedding more light on the critical role of product innovation.

The nature of axiom 1 is common sense, which is demonstrated clearly by examples and reasoning provided. That is why the author calls it an axiom. There is no need of empirical evidence for common sense or for an axiom (e.g. the straight line is the shortest between two points), but economists require empirical evidence for everything. The direct proof for axiom 1 would be an experiment on anybody using any goods or services, such as those showed in our hypothetical examples like ice cream, shoes, music, and massage. If you do not agree with these common-sense examples, you can do experiments on yourself, e.g. listening to music 24 hours on end.

The indirect proof at a large scale can be found in product cycle theory, market saturation phenomenon, and the backward Engel curve for inferior goods. However, using this kind of indirect proof need to be accompanied by careful logical reasoning. For example, Engel curve is obtained by expenditure data on different commodity groups by different household income groups, so this bears some similarity to the effect of increasing income on an individual, but they are not exactly the same thing (the collective behaviour of households of different income groups may not be the same as that of one individual when income rises). Moreover, the data is based on commodity groups and income groups, which gives the empirical results an indicative or gross nature. A commodity group like 'computer' includes computers of many different brands, speeds, capacity, etc. A person would only need one desktop and one laptop (say this is the satiation point), but the advent of new high speed computer may induce him buy one to replace the old one. This is an example that innovation changes satiation point of a commodity group but does not change the satiation point of a specific commodity (e.g. one needs only one computer of the same brand and same model). The income group has the similar issue and, more importantly the expenditure data may not be able to include sufficiently high income levels (i.e. 'out of sample') at which the satiation point appears. As a result, despite the fact that most goods do not have a backward Engel curve, and that Chai and Moneta (2014) showed that the Engel curve can change position and shape over time, these facts do not prove that a common sense like Axiom 1 is incorrect.

Axiom 2 (dual role of savings): savings act both as precautionary premium and as saved resources.

Saving is traditionally treated as future consumption plus some interest income from the saving, so it is normally not included in a utility function. The theory behind this practice is the life cycle theory developed by Modigliani and Brumberg (1954) and Modigliani (1986), and the permanent income theory developed by Friedman (1957). Both theories utilized the intertemporal choice model developed by Fisher (1930). However, these studies assumed away any uncertainty in our life and thus ignored the fundamental function of saving – saving for rainy days, i.e. precautionary saving.

Many researchers have studied precautionary saving (e.g. Leland, 1968, Kimball, 1990, Weil, 1993). Since it is related to uncertainty, the precautionary saving in the studies is treated as an insurance premium and is determined by intertemporal utility maximization based on the expected utility function. In the current paper, the author argues that the nature of precautionary saving indicates that savings can generate utility directly and immediately – with savings in hand the person at once feels more secure and thus happier.

In fact, in considering the decision on how much to consume and to save today is a choice under uncertainty, savings play a dual role. One is that savings can be viewed as premium paid to reduce the uncertainty, namely the equivalent precautionary premium in Kimball (1990). The other role is that, as usual, savings are saved resources to be used in the future. The utility of savings as precautionary premium is achieved when the saving action takes place, so this utility is in current period and thus savings have to be put into the utility function in the current period. The utility of savings as saved resources is achieved when the savings are consumed in the future period, so this part of utility of savings should be expressed as utility of increased consumption in the future period.

Including savings in a utility function is not a novel practice either. For example, Howe (1975) treated saving as a good in the current period in a linear expenditure system. In so doing, Howe derived the same extended linear expenditure system as developed by Lluch (1973), who used an intertemporal utility maximization of the Stone-Geary utility function. However, it is worth noting that the different consideration of putting savings into the utility function in Howe (1975) and in this study. In Howe (1975), the savings in the utility function means saved resources or equivalent of future consumption, so the uncertainty of future is completely ignored. In this study, the savings in the utility function means precautionary premium to reduce the uncertainty in the future.

The amount of utility generated by the role of savings as precautionary premium at the present time is positively related to the amount of savings, i.e. the more one saves, the more secure one feels. Since there is no absolute security for any individual, there is no limit on the security feeling from savings action. As a result, there is no ceiling on the intention to save and the concept of ‘the more is the better’ applies here with no constraint. A number of utility functions can be used to describe these nature of saving action. For simplification, we use a linear function for the utility of saving: $U(Savings) = \alpha_S * Savings$. This utility function of saving also allows negative saving (dissaving) to generate negative utility. In considering all factors affecting household utility, the following utility function can be created:

$$U = U(x_1, x_{2,} ... x_n, Savings) = \sum_{i=1}^{n} \alpha_i (2m_i x_i - x_i^2) + \alpha_S * Savings$$

α_S is the weight of savings in the individual’s utility.

There may be some arguments against putting saving into a utility function. One such argument is that, if savings generate utility and dis-savings (net borrowing) generate negative utility, how can you explain why so many households take on debts (e.g. housing mortgage loan)? People raising this question only see the one type of utility related to savings – the utility generated from saving action. In the case of household debt (dissaving), the current utility related to the action of dissaving is negative, but the dissaving allows the household to consume more in this period and this generates more utility from future consumption. As long as the future utility from the increased consumption outweighs the negative current utility from the dissaving action, the total intertemporal utility increases when the household borrows. As a result, the rising household debt phenomenon does not disprove the validity of putting saving into a utility function.

Axiom 3 (investment-consumption dependency): to be profitable, investment must be used to cater for future consumption.

Investment demand is usually treated as an independent final demand, which has the same status as consumption, but Axiom 3 proposes that investment demand is not an independent one. The logic for investment-consumption dependency is quite simple. The purpose of investment is to make profits. This aim can be achieved only when the products resulting from the investment can be sold to consumers. In other words, to be profitable, investment must be used to cater for future consumption. Consequently, investment today is dependent on the expected consumption growth in the future, or consumption growth potential. This simple logic can be argued in a more rigorous manner.

What determines investment demand? Tobin (1969) noticed that investment demand depends critically on profitability. A further thinking reveals that the profitability is in turn dependent crucially on the final demand.

We think about the case of investment in production first. In this case, the profitability of an investment is determined by both cost of, and revenue from, an increase in production capital. On the cost side, interest rates (or capital rental price), taxation policy, wages and prices of intermediate inputs are important factors, we can use interest rate as a representative of investment cost. On the revenue side, investment income is achieved through sales of output to both intermediate demand and final demand. Since the sales to intermediate demand is to cater for the final demand, the final demand will be crucial in determining both price and the quantity of sales made.

Similar reasoning can be applied to investment in assets, such as housing, bonds, and shares. On the cost side, interest rates (or borrowing rates) are a significant factor. On the revenue side, the sales of assets ultimately depend on the final demand for products. For example, if you invest in company shares or in housing assets, the revenue is determined by the prices of shares or housing prices, which are ultimately determined by the sales of the company's products or by the renting and/or selling of the housing to the final demand.

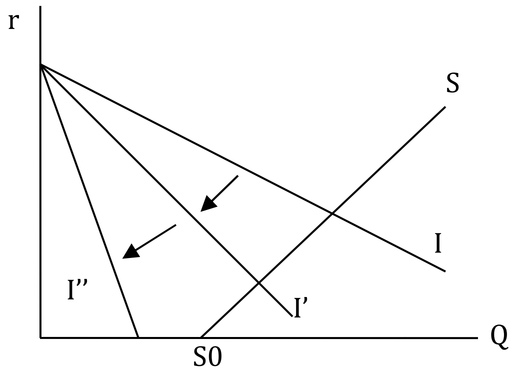
There are many kinds of final demand such as household consumption, government consumption, exports, and investment demand. Household consumption, government consumption, and exports simply reflect consumption by different consumer groups, so they can be grouped under a broader definition of household consumption for the purpose of simplicity. As such, there are two types of final demands: the investment demand and the broadly defined household consumption. Since our aim is to discover what determines investment demand and, apparently, investment demand cannot be a determinant of itself, the fundamental final demand – household consumption – must be the determinant of investment demand.

The direct link between investment and consumption did not draw much attention from economists. Keynesian economists blamed the lack of investment demand on the lack of animal spirits and on liquidity preference. The traditional investment theory (e.g. the loanable funds theory) proposed by classical economists claims that a flexible interest rate can *always* make sure that the supply of savings is equal to the demand for investment, thereby producing equilibrium in the capital market. Keynesian economists rejected this theory based on liquidity preference, which rejects the neutrality of money, the core assumption of neoclassical economics.

Actually, even with the assumption of neutrality of money, i.e., without liquidity preference, the loanable funds theory does not hold. This can be shown in Figure 2. Because of precautionary saving, in a general case (i.e. a case does not rely on accumulated savings from the previous years because the household has to maintain a fixed balance sheet in the long run), there are always some positive savings in an economy (S0 shown in Figure 2) even if the interest rate is zero. If the investment decreases dramatically (left shift of investment curve from I to I″ in Figure 2), there will be excessive savings although the interest rate is perfectly flexible.

It should be pointed out that Figure 2 shows only the simplified case but this does not lose the generality. For example, it can be argued that, in the case of having a stock of liquid wealth, the flow of savings can be zero or even negative (dissaving). The first response to this argument is that the positive stock of liquid wealth indicates that the saving flow in the past is positive. Secondly, relying on past saving stock cannot not last forever, so this is a special case in short run and thus does not invalidate the possibility of uninvested savings in the long run. From this point of view, we must keep the balance sheet unchanged to present a general case.

Figure 2. *The possibility of uninvested savings*



Some readers may argue that the saving cannot finance the investment in the same period so the saving-investment balance in Figure 2 is mispresented. Figure 2 does not show time frame in order to simplify the presentation. Considering different periods, the investment in Figure 2 can be viewed as the perspective investment in period 1, which is equal to the investment in period 2, so the investment can balance the savings in period 1. This reasoning is the same as that for the supply/demand schedule in the goods market.

Goods must be produced first and then can be consumed, so this involves two periods. We can used the perspective demand to overcome this time difference and produce a supply/demand equilibrium.

One may also argue that although the interest rate is determined by investment and savings in Figure 2, in reality the interest rate is managed by reserve bank through the control of money supply. This argument mixed real interest rate with nominal interest rate. The interest rate determined in Figure 2 is the real interest rate while the interest rate stipulated by the reserve bank is nominal interest rate. The Reserve bank generally adopts a targeting-inflation interest rate policy, i.e. letting nominal interest rate keep up with the inflation rate. This will lead the real interest rate unchanged because real interest rate = nominal interest rate – inflation rate. Given the unchanged real interest rate, investment decision will not be affected. In short, it is the real interest rate, rather than the nominal interest rate, that are relevant to the decision on investment and savings.

More generally speaking, investment demand is determined by both the cost side and the revenue side of investment. On the cost side, interest rate is a good indicator. On the revenue side, the broadly-defined household consumption, or more accurately, the household consumption growth potential determines investment profitability. The potential of household consumption growth can be represented by the difference between the consumption ceiling, and the current consumption level. As such, we can have the aggregate investment demand function for the economy:

$$I = \frac{B}{1+r}\left(\sum_{i=1}^{n} m_i w_i - \sum_{i=1}^{n} c_i w_i\right)$$

Where B is a parameter indicating the propensity to invest when no borrowing cost exists; r is interest rate, which discounts B; w_i is the weighting parameter aggregation.

The above investment function is very different from the traditional investment function: there is no role for income. In a conventional macroeconomics, investment demand is positively related to income. The reasoning is that higher income may indicate higher future consumption (again, future consumption is the key for investment). However, this reasoning is problematic during a recession or in considering consumption satiation: higher income enables high purchasing power but not necessarily leads to high consumption. To avoid this problem, the investment function in this paper links the future consumption to consumption growth potential. The impact of income on investment in this paper is indirectly reflected on savings, which ration the investment. The impact of income on savings has been taken care of by including savings in the utility function.

One may argue that the 3 axioms are essentially assumptions or hypotheses imposed by the author. What is the empirical evidence for them? The author’s response is that the axioms in this paper are logical conclusion based on common sense. Since everyone has a limit in consuming a good in limited time period, the axiom 1 – consumption ceiling – is obvious. If one agrees that savings generates both security satisfaction at the present time and consumption satisfaction in the future, the axiom 2 – dual utility of savings – is easily comprehended. Because the purpose of investment – to make profits – can be achieved only through an increase in future consumption of goods induced by investment, the axiom 3 – the investment consumption dependency – is a natural and logical conclusion. We can find many trivial cases to supports our common sense, for example, the consumption of ice cream shown in explaining axiom 1, but it is not necessary to pile up the trivial examples to support common sense.

Among the 3 axioms in this section, axiom 1 (the consumption ceiling) plays a vital role and becomes the fundamental cause for economic stagnation or recessions. The lack of new products due to the slow speed of product innovation necessitates consumption ceilings. As the gap between the consumption ceiling and the actual consumption level narrows, consumption growth reduces. This also leads to a reduction in investment. As demand falls short of production capacity, not all factors of production (i.e. labour and capital) will be utilized and thus the economy enters stagnation or a recession. Facing the stagnation of sales, firms have to put more effort in product innovation. New products eventually appear and the consumption ceiling is lifted. Consumption and investment increases and the economy recovers and expands. As a result, the only way to avoid a recession is to stimulate product innovation so as to lift consumption ceilings. These are intuitions from the 3 axioms. Next we will employ a multi-commodity economic model to demonstrate more rigorously how and when an economic recession occurs.

4. The multi-commodity macroeconomic model

The existing theories regarding economic growth and business cycles are based on aggregate macro models (e.g. Keynes' multiplier model, the AS/AD model, the life-cycle permanent-income model, the real business cycle model and the endogenous growth model). An aggregate macro model has only one good for the consumer (i.e. consumption) so the model is unable to capture market dynamics and market saturation. On the other hand, there are a number of multi-commodity models, for example product variety models and numerous empirical computable general equilibrium models. However, the household utility functions used in these models are normally CES function or LES function. In these functions, there is no limit on consumption of any goods, so the model cannot reflect consumption situation or market saturation phenomenon. To demonstrate mathematically how the three axioms proposed in the previous section can lead to an economic stagnation and thus to uncover the genesis of economic recessions, this section adopts a general equilibrium approach to build a multi-commodity macro model, with the three axioms embedded in the consumption function and investment function in the model. To sharpen the focus, the multi-commodity model used in this section is a static one, but a dynamic upgrade is provided for interested readers.

4.1. The static model

The economy in the model consists of one representative household and n representative firms. For simplicity, government is not included in the model but the function of government is implicitly included in the broad definition of households. Government spending and investment is similar to that for households. The function of government taxation and social welfare influences income distribution, which is reflected by an income distribution parameter in the model. Since income inequality is not the focus of the study, only one representative household is used in the model. This means that income inequality as well as lending and borrowing are not explicitly considered in the model[(2)]. However, they are indirectly included in the income distribution parameter – lending and borrowing can lead to temporarily more equitable income distribution. Also for simplicity, a closed economy is assumed, so international trade and finance are not included in the model.

The basic transactions in the model are as follows. The household provides labour and capital to all firms, and obtains wages and capital rentals in return. The household also uses its income to purchase goods from firms for consumption purposes, and supplies its savings to firms for investment purposes. Under the zero economic profit condition, each firm uses labour, capital and technology to produce one unique commodity for the economy, and decides on its requirements for labour, capital, and investment in production.

4.1.1. Household consumption and savings

We consider the household first. The ultimate goal of a society is to maximize household utility (other goals such as investment, accumulation and development are parts of household utility in the future). This means that household utility is a crucial part of an economy-wide model. The utility function described in the previous section requires further modification before it is used in the model. First, since commodity demand includes both consumption demand and investment demand, we use 'c_i' to replace 'x_i' in the utility function in the previous section to explicitly indicate consumption demand. Second, we need to consider the fact that there are a large number of households in an economy and that the distributional effect is an important factor in household consumption and utility. It is desirable to develop a multi-household model to include the distributional effect. This would, however, complicate the model and thus interfere with the main purpose of the paper. Instead, the author adds a distributional effect parameter in the utility function of the representative household. Finally, the varieties of commodities may increase due to product innovation, so n+Δn is used to reflect this effect. For simplicity, this paper does not model the determinants of product innovation, so n+Δn is left as exogenous. The new utility function is as follows:

$$U = U(c_1, c_{2,} ...c_{n+\Delta n}, Savings) = \sum_{i=1}^{n+\Delta n} \alpha_i (2\theta m_i c_i - c_i^{\ 2}) + \alpha_S * Savings \qquad (1)$$

Where θ is the distributional parameter, $0<\theta\leq 1$. θ =1 indicates that every household in the household group has the same level of income. When income distribution is not equal, some households cannot reach their consumption saturation point due to lack of income support. In other words, their consumption ceilings are practically lowered due to income constraint. This effect is captured by making $\theta<1$.

The budget can be expressed as:

$$Y = \sum_{i=1}^{n+\Delta n} P_i * c_i + \sum_{i=1}^{n+\Delta n} P_i * S_i$$

Where, c_i is consumption of each commodity, S_i is the saved of each commodities, P_i is the commodity price.

To obtain aggregate real savings, we need a weighting parameter for aggregation. Letting it be w_i, we have

Savings = $\sum w_i \times S_i$.

Defining δ_i as the share of each commodity saved (S_i) in total savings, i.e. $\delta_i = S_i$ / Savings, we can obtain the price for aggregate savings:

$$P_S = \left(\sum_{i=1}^{n+\Delta n} P_i * S_i \right) / \, Saving = \sum_{i=1}^{n+\Delta n} P_i * \delta_i .$$

As such, the optimal consumption problem for households can be expressed as:

Maximize $U = U(c_1, c_{2,} ... c_{n+\Delta n}, Savings) = \sum_{i=1}^{n+\Delta n} \alpha_i (2\theta m_i c_i - c_i^{\,2}) + \alpha_S * Savings$

Subject to $Y = \sum_{i=1}^{n+\Delta n} P_i * c_i + P_S * Savings$

Setting up a Lagrangian expression:

$$1 = U + \lambda(Y - \sum_{i=1}^{n+\Delta n} P_i * c_i - P_S * Savings)$$

Using the first order condition we can derive the optimal consumption of good i as follows:

$$c_i = \theta m_i - \frac{\alpha_S}{2\alpha_i} \frac{P_i}{P_s} \tag{2}$$

$$Savings = (Y - \sum_{i=1}^{n+\Delta n} \theta P_i m_i + \sum_{i=1}^{n+\Delta n} \frac{P_i^2 \alpha_S}{2\alpha_i P_S}) / P_S \tag{3}$$

4.1.2. Firm's investment and unsold stock

Since the firm's investment decision has an impact on its production, we discuss the firm's investment first. It is assumed that the firm can identify consumption ceilings as well as the impact of the distributional effect on consumption, so the firm can invest a proportion of consumption growth potential, i.e. the gap between the constrained consumption ceiling and the current consumption. Letting the investment demand for commodity i (I_i) is proportionally related to consumption growth potential and to the propensity to invest after being discounted by interest rate, we have the following investment demand function for each commodity:

$$I_i = \frac{B}{1+r}(\theta m_i - c_i)$$

Where B indicates the propensity to invest, r is interest rate, m_i is the maximum amount of consumption on good i, c_i is the actual amount of consumption of good i.

It assumed that $0 \leq B \leq 1$. When B=1, the firm invests the highest amount in production to produce a maximum amount of goods which will be purchased by the household.

To obtain the aggregate real investment, we need a weighting parameter for aggregation. We use the same weighting w_i as that for savings because, in the case of the existence of market clearance, $I_i=S_i$ and I=Saving. The same weighting ensures the amounts of investment and saving at both aggregate and disaggregate levels to be consistent with market clearance. As such, we have

$I = \sum w_i \times I_i$

$$I = \frac{B}{1+r}\left(\sum_{i=1}^{n+\Delta n} \theta m_i w_i - \sum_{i=1}^{n+\Delta n} c_i w_i\right) \tag{4}$$

With the investment functions at both disaggregated and aggregate levels, we can calculate the share of each commodity in total investment: $\beta_i=I_i/I$

It is worth mentioning that, for simplicity, we use the same parameter B for the investment for all commodity so we have the same overall propensity to invest in the total investment demand function. This treatment does not lose generality. If one uses different parameters as propensity of investment demand for different commodities, the parameter for propensity of overall investment demand will be the weighted average of the parameters for all commodities. The only difference is that calculation of weight average is required to obtain the parameter for overall propensity to invest.

This investment demand is financed by household savings. The uninvested household savings (the gap between saving and investment demand) equals the unsold stock (S) or inventory at firm (because the firm has paid the household of the value inventory but failed to sell it), so we have.

$S = Savings - I = \sum w_i S_i - \sum w_i I_i$

In a static model for a closed economy, the unsold stock S in the above equation should be non-negative because investment demand must be financed by savings. However, the unsold stock S can be negative when the economy is open or when the model has multiple periods. The additional finance in this case can come from overseas or from past savings. In considering the accumulated past savings (e.g. wealth), the savings in a dynamic model can be negative, i.e. dissaving.

4.1.3. Firm's input demand

To depict the firm's production, the following Cobb-Douglas function is used for the purpose of simplicity:

$$x_i = (A_i + \Delta A_i) * L_i^{\gamma_i} * K_i^{1-\gamma_i}$$

Where L means labour, K mean capital, A indicates the level of technology, and γ is the share of labour in total inputs.

The optimal production problem can be expressed as:

Minimize $Cost = P_L * L_i + P_K * K_i$

Subject to $Output = x_i = (A_i + \Delta A_i) * L_i^{\gamma_i} * K_i^{1-\gamma_i}$

Setting up a Lagrangian expression:

$$1 = P_L * L_i + P_K * K_i + \lambda [x_i - (A_i + \Delta A_i) * L_i^{\gamma_i} * K_i^{1-\gamma_i}]$$

Using the first order condition we can show optimal demand for labour and capital as follows:

$$L_i = \left(\frac{x_i}{A_i + \Delta A_i} \right) \left(\frac{\gamma_i P_K}{(1-\gamma_i) P_L} \right)^{1-\gamma_i} \tag{5}$$

and

$$K_i = \left(\frac{x_i}{A_i + \Delta A_i} \right) \left(\frac{(1-\gamma_i) P_L}{\gamma_i P_K} \right)^{\gamma_i} \tag{6}$$

These results link the firm's demand for labour and capital to the firm's output x_i. More generally, the results show that the factor market is closely related to the commodity market.

4.1.4. Resource constraint and market clearance condition

The resources constraint in a closed economy can be expressed as

Savings is not less than investment: $S = Savings - I = \sum w_i S_i - \sum w_i I_i \geq 0$

Labour supply is not less than labour demand: $L \geq \sum L_i$

Capital supply is not less than capital demand: $K \geq \sum K_i$.

Finally, we consider the market clearance condition. The total supply of commodity x_i in the economy is the sum of both the consumed and unconsumed commodity, namely, $x_{Si} = c_i + S_i$. On the other hand, the total demand for commodity x_i comprises the consumption demand and the investment demand, so that the total demand for x_i can be expressed as $x_{Di} = c_i + I_i$.

Thus, the excess demand function for x_i is: $ED_i = x_{Di} - x_{Si} = I_i - S_i$.

The conditions for market clearance is that, for each commodity i, $ED_i = x_{Di} - x_{Si} = I_i - S_i$ = 0, or I_i = S_i. Aggregating all commodities, we have market clearance condition: $\sum w_i I_i = \sum w_i S_i$, or I = Savings.

In a general equilibrium model, investment is always equal to savings because it is simplistically and idealistically assumed that all savings are invested in neoclassical economics, so I=Savings is guaranteed at all time. With this guarantee, a general equilibrium is achievable at any time: if $I_i \neq S_i$ for some or all commodity types, the price mechanism will kick in and adjust any difference between investment and savings in all commodity types.

However, in our static model, investment is determined by consumption growth potential (the difference between current consumption and the maximum consumption) while saving is determined by utility maximization procedure, so there is no guarantee that total investment equals total savings – only the resource constraint (Savings ≥ I) is applied to savings and investment. As a result, the general equilibrium is not guaranteed: if Savings > I, the price mechanism cannot work out solution for $S_i = I_i$ because there is an overall oversupply in the economy.

4.2. The recursive dynamic model

The model presented so far is a static one, which can be easily upgraded by adding time frame and by considering dynamics in technology, capital, and wealth.

Since wealth is accumulated savings, so the wealth dynamic can be described by equalling W_{t+1} (wealth in time t+1) to W_t (wealth in time t) plus savings in time t.

$W_{t+1} = W_t + Savings_t$.

The technological and capital dynamics for industry i can be expressed as

$A_{i,\,t+1} = A_{i,\,t} + \Delta A_{i,\,t}$, where $\Delta A_{i,\,t}$ is the change of technology in industry i at time t.

$K_{i,\,t+1} = (1\text{-}\delta)K_{i,\,t} + I_{i,\,t}$, where $I_{i,\,t}$ is the change of technology in industry i at time t, δ is the depreciation rate.

The household utility function at time t can be written as:

$$U_t = U(c_{1,t}, c_{2,t} \dots c_{n+\Delta n,t}, Savings_t) = \sum_{i=1}^{n+\Delta n} \alpha_{i,t}(2\theta_t\, m_{i,t}c_{i,t} - c_{i,t}{}^2) + \alpha_{S,t} * Savings_t$$

This utility function is subject to a budget: $W_t + Y_t \geq \sum P_{i,t}\, c_{i,t} + P_{S,t} \sum S_{i,t}$.

The production function and investment demand function are as follows:

$x_{i,t} = A_{i,t} * L_{i,t}^{\gamma_i} * K_{i,t}^{1-\gamma_i}$, subject to production cost = $P_{L,\,t} \times L_{i,\,t} + P_{K,\,t} \times K_{i,\,t}$

$$I_t = \frac{B_t}{1+r_t}\left(\sum_{i=1}^{n+\Delta n} \theta_t m_{i,t} w_{i,t} - \sum_{i=1}^{n+\Delta n} c_{i,t} w_{i,t}\right) \tag{7}$$

In a similar fashion to the static model, the household uninvested savings equals the unsold stock (i.e. inventory) at firm, which are calculated as the difference between savings and investment demand,

$S_t = Savings_t - I_t$

Investment demand is financed by savings; any uninvested savings (or stocks) will be accumulated as wealth; and any excess investment demand over savings will be drawn from wealth.

4.3. The intertemporal equilibrium dynamic model

The above equations transforms the static model to a recursive dynamic model. This recursive model can generate the equilibrium for each period. The results from previous period will have an influence on the results in the next periods. For example, I_t (the investment in time t) will affect K_{t+1} (the capital in time t+1), which in turn will affect a number of variable in time t+1 such as output level (x_{t+1}), savings (S_{t+1}), consumption (c_{t+1}), and investment (I_{t+1}). However, the mechanism determining the equilibrium or disequilibrium in each period is the same as in the static model.

The above recursive model is unable to determine either the intertemporal equilibrium or the optimal time path for the economy, so an intertemporal model is needed. Since an intertemporal equilibrium model with multi-commodity is very complex, we have to be content with a Ramsey/Solow-style one-commodity model, but the essence of including multi-commodity in the static model – consumption ceiling – will be reflected in the intertemporal model. Moreover, to reduce the number of variables, we follow the common practice to eliminate the variable labour by measuring capital, output, utility, and consumption in per capita term. In a traditional Ramsey/Solow model, all savings are assumed to be invested. This assumption is implausible thus has to be relaxed. Moreover, the three axioms featured in the static model will also be used in the intertemporal equilibrium model.

Since we are considering an intertemporal equilibrium model, we have to use continuous time, which is different from the discrete time in the recursive model. Also, because all variables in the intertemporal model is measured in per capita term, we use the lower case for most variables so as to differentiate them from aggregate variables, e.g. using k for capital, c for consumption, and s for savings.

In per capita term, the function can be written as:

$y = a \times f(k)$, where a is technology, k is capital per worker, and $f(k)' > 0$.

The household utility function in per capita term can be written as:

$U(c) = \alpha_C \times (2mc - c^2) + \alpha_S \times (af(k) - c)$, where, α_C is the weighting for utility from consumption, α_S is the weighting for utility from saving, m is consumption ceiling, c is actual consumption, $c \leq m$.

The investment per capita is proportional to the gap between the consumption ceiling and the actual consumption level, so it can be written as

$I = b(m - c)$, where I is investment and b is an interest-rate-discounted investment propensity parameter, $b = B / (1 + r)$, $0 < b < 1$; as before, B is the propensity to invest, r is the interest rate.

Based on this investment demand function, the per capita capital dynamics can be written as:

$k' = \Delta k = b(m - c) - \delta k - nk$, where δ is the capital depreciation rate and n is the growth rate of population (or labour force).

The dynamics of per capita assets is determined by uninvested household saving or unsold stock (inventory) at the firm:

$s' = \Delta s = \text{Savings} = y - c - i = af(k) - c - b(m - c)$, where s stands for stock.

Considering a time preference rate (or future discount rate) of θ, the optimal control problem can be expressed as:

Maximize $V = \int_0^{\infty} U(c)e^{-\theta t}dt$

Subject to $s' = af(k) - c - b(m - c)$, $k' = b(m - c) - \delta k - nk$, $s(0) = 0$, $s(\infty) \geq 0$, $k(0) = 0$, $k(\infty) \geq 0$, and $0 \leq c(t) \leq m$.

The standard Hamiltonian function is

$H = U(c)e^{-\theta t} + \lambda_1[af(k) - c - b(m - c)] + \lambda_2[b(m - c) - \delta k - nk]$.

From this Hamiltonian function it is easy to see that, if we impose a condition that $\lambda_1 = \lambda_2$, then the model collapses to the traditional Ramsey/Solow model.

The current-value form of Hamiltonian function is

$H_c = U(c) + \eta_1[af(k) - c - b(m-c)] + \eta_2[b(m - c) - \delta k - nk]$,

where $\eta_1 = \lambda_1 e^{\theta t}$, and $\eta_2 = \lambda_2 e^{\theta t}$.

The necessary condition for an optimal solution is that, at each t,

(i) $\partial H_c/\partial c = 0$

(ii) $\eta_1' = -\partial H_c/\partial s + \theta\eta_1$

(iii) $\eta_2' = -\partial H_c/\partial k + \theta\eta_2$

(iv) $s' = \partial H_c/\partial \eta_1$

(v) $k' = \partial H_c/\partial \eta_2$

(vi) $s(0) = 0$, $s(\infty) \geq 0$, $k(0) = 0$, $k(\infty) \geq 0$, $0 \leq c(t) \leq m$.

Condition (ii) gives: $\eta_1' = -\partial H_c/\partial s + \theta\eta_1 = \theta\eta_1$. The obvious solution for η_1 is $\eta_1 = e^{\theta t}$. Condition (i) gives $u' + \eta_1(-1 + b) + \eta_2(-b) = 0$, or $\eta_2 = \eta_1(b - 1)/b + u'/b$.

Considering $\eta_1 = e^{\theta t}$ and $u = \alpha_C \times (2mc - c^2) + \alpha_S \times (af(k) - c)$, we have $\eta_2 = e^{\theta t}(b - 1) / b + (2\alpha_C m - 2\alpha_C c - \alpha_S) / b$, and thus $\eta_2' = \theta e^{\theta t}(b - 1)/b + 2\alpha_C c'/ b$.

Condition (iii) gives:

$\eta_2' = -\partial H_c / \partial k + \theta\eta_2 = -[\eta_1 af(k)' - \eta_2(\delta + n)] + \theta\eta_2 = -\eta_1 af(k)' + \eta_2(\delta + n + \theta)] = -e^{\theta t}af(k)' + [e^{\theta t}(b - 1) / b + (2\alpha_C \times m - 2\alpha_C \times c - \alpha_S) / b](\delta + n + \theta)$.

Based on above two equations, we have: $\eta_2' = \theta e^{\theta t}(b - 1) / b + 2\alpha_C \times c' / b = -e^{\theta t}af(k)' + [e^{\theta t}(b - 1) / b + (2\alpha_C \times m - 2\alpha_C \times c - \alpha_S) / b](\delta + n + \theta)]$, so,

$c' = -0.5\theta e^{\theta t}(b - 1) - 0.5e^{\theta t}abf(k)' + 0.5[e^{\theta t}(b - 1) + (2\alpha_C \times m - 2\alpha_C \times c - \alpha_S)](\delta + n + \theta)$,

$c' = -0.5e^{\theta t}[abf(k)' + (1 - b)(\delta + n)] + (\alpha_C \times m - \alpha_C \times c - 0.5\alpha_S)(\delta + n + \theta)$ (8)

Conditions (iv) and (v) gives the growth rate of stock and capital respectively:

$s' = af(k) - c - b(m - c)$ (9)

$k' = b(m - c) - \delta k - nk.$ (10)

These conditions will be used in the next section.

5. Results interpretation

Using the above model and the concept of excess demand (or excess supply), we can discuss how and when an economic recession will occur, as well as its features. The results from the dynamic version are more complex but have features similar to the static version: the consumption ceiling will eventually cause economic stagnation. The section discusses the results from three perspectives: the demand-side perspective, the supply-side perspective, and the dynamic perspective.

5.1. Results from the static model: a demand-side perspective

In this section, we assume the prices in the static model is fixed so that we can derive some intuitive but essential results from the model. For an economy to grow without a recession, the condition for general equilibrium must be satisfied, i.e. the total supply of a commodity must be cleared by the market. The excess demand function for x_i is: $ED_i = x_{Di} - x_{Si} = I_i - S_i$. Aggregating the excess demand functions for all commodities, we arrive at $ED = \sum_{i=1}^{n+\Delta n} I_i w_i - \sum_{i=1}^{n+\Delta n} S_i w_i = I - Saving$.

This equation indicates that the equilibrium of the commodity market hinges on the balance of saving and investment. If total investment demand is greater than total savings, there will be overall excess demand for commodities in the economy. On the other hand, if savings cannot be fully used for investment purposes, there will be an overall excess supply of commodities. To achieve the market clearance and thus to avoid an economic recession, the overall excess demand ED must be non-negative.

Recalling the investment equation (equation 4), we can express the condition to avoid a recession as:

$$ED = \frac{B}{1+r}\sum_{i=1}^{n+\Delta n}\theta m_i w_i - \frac{B}{1+r}\sum_{i=1}^{n+\Delta n} c_i w_i - Saving \geq 0$$

Plugging the consumption equation (equation 2) and the saving equation (equation 3) into the above inequality, we have:

$$\frac{B}{1+r}\sum_{i=1}^{n+\Delta n}\theta m_i w_i - \frac{B}{1+r}\sum_{i=1}^{n+\Delta n}(\theta m_i w_i - \frac{\alpha_S P_i w_i}{2\alpha_i P_S}) - (Y - \sum_{i=1}^{n+\Delta n}\theta P_i m_i + \sum_{i=1}^{n+\Delta n}\frac{P_i^2 \alpha_S}{2\alpha_i P_S}) / P_S \geq 0$$

$$Y \leq \sum_{i=1}^{n+\Delta n}\theta P_i m_i + \sum_{i=1}^{n+\Delta n}\frac{BP_i \alpha_S w_i}{2(1+r)\alpha_i} - \sum_{i=1}^{n+\Delta n}\frac{P_i^2 \alpha_S}{2\alpha_i P_S} \tag{11}$$

This inequality shows that, to avoid a recession, the household income must be below certain level! To allow income to increase without a ceiling, one may increase θ (i.e. improving the equality in income distribution) or increase B (the propensity to invest) or decrease r (the interest rate), but the effect of these efforts is limited because the maximum value of both θ and B is 1 and the minimum value of r is 0. The only way to allow the income level to increase unrestrictedly is to increase Δn, i.e. inventing new products. Since $\theta \times m_i$ is much larger compared with $P_i\alpha_s/2\alpha_i P_s$ (m_i is generally very high compared with other items here), when Δn increases, the increase in the first term on the right hand side will outweigh the increase in the third term and thus the cap on Y will be lifted.

Household supply of labour and capital is determined by household willingness to obtain income, which in turn is determined by consumption and savings. So, the household will supply the amount of labour and capital to produce the amount of output of good i equal to the sum of the consumed and the saved by the household. In this reasoning, we substitute $x_i = c_i + S_i$ into equations (5) and (6) and obtain the amount of labour and capital supplied for the production of good x_i:

$$L_{Si} = \left(\frac{c_i + S_i}{A_i + \Delta A_i}\right)\left(\frac{\gamma_i P_K}{(1-\gamma_i)P_L}\right)^{1-\gamma_i} \tag{12}$$

$$K_{Si} = \left(\frac{c_i + S_i}{A_i + \Delta A_i}\right)\left(\frac{(1-\gamma_i)P_L}{\gamma_i P_K}\right)^{\gamma_i} \tag{13}$$

On the other hand, demand for labour and capital is determined by final demand $c_i + I_i$. Therefore, including $x_i = c_i + I_i$ into equations (5) and (6) we have labour and capital demand functions:

$$L_{Di} = \left(\frac{c_i + I_i}{A_i + \Delta A_i}\right)\left(\frac{\gamma_i P_K}{(1-\gamma_i)P_L}\right)^{1-\gamma_i} \tag{14}$$

$$K_{Di} = \left(\frac{c_i + I_i}{A_i + \Delta A_i} \right) \left(\frac{(1-\gamma_i)P_L}{\gamma_i P_K} \right)^{\gamma_i} \tag{15}$$

Excess demand in the factor market is the sum of the excess demand for labour and capital in producing each commodity, namely:

$$ED_L = \sum_{i=1}^{n+\Delta n} ED_{Li} = \sum_{i=1}^{n+\Delta n} L_{Di} - \sum_{i=1}^{n+\Delta n} L_{Si} \tag{16}$$

$$ED_K = \sum_{i=1}^{n+\Delta n} ED_{Ki} = \sum_{i=1}^{n+\Delta n} K_{Di} - \sum_{i=1}^{n+\Delta n} K_{Si} \tag{17}$$

Substituting equations (12) to (15) into the above equations and utilizing the saving share δ_i and investment share β_i, we have:

$$ED_L = \sum_{i=1}^{n+\Delta n} \left(\frac{I_i - S_i}{A_i} \right) \left(\frac{\gamma_i P_K}{(1-\gamma_i)P_L} \right)^{1-\gamma_i} = \sum_{i=1}^{n+\Delta n} \mu_i I_i - \sum_{i=1}^{n+\Delta n} \mu_i S_i = \sum_{i=1}^{n+\Delta n} \mu_i \beta_i I - \sum_{i=1}^{n+\Delta n} \mu_i \delta_i S \tag{18}$$

$$ED_K = \sum_{i=1}^{n+\Delta n} \left(\frac{I_i - S_i}{A_i} \right) \left(\frac{(1-\gamma_i)P_L}{\gamma_i P_K} \right)^{\gamma_i} = \sum_{i=1}^{n+\Delta n} v_i I_i - \sum_{i=1}^{n+\Delta n} v_i S_i = \sum_{i=1}^{n+\Delta n} v_i \beta_i I - \sum_{i=1}^{n+\Delta n} v_i \delta_i S \tag{19}$$

where

$$\mu_i = \left(\frac{\gamma_i P_K}{(1-\gamma_i)P_L} \right)^{1-\gamma_i} / (A_i + \Delta A_i), \quad v_i = \left(\frac{(1-\gamma_i)P_L}{\gamma_i P_K} \right)^{\gamma_i} / (A_i + \Delta A_i)$$

The above two equations show that the excess demand for both labour and capital is the difference between weighted total investment and weighted total savings. This is very similar to the excess demand function in the commodity market. Thus, if the investment demand is greater (smaller) than savings, there is an excess demand (supply) in commodity markets, there will be an excess demand for (supply of) labour and capital. The size of excess demand (supply) in different markets will, however, differ due to the weights. Since demand for primary factors closely links to demand for commodities, the reasons for excess supply in the commodity market can also explain the excess supply in the factor market.

5.2. Results from the static model: a supply-side perspective

The above demand-side approach gives us an intuitive picture on commodity and factor markets. However, this picture is not a high-resolution one because the prices for the commodities and for factors are set as exogenous. In fact the prices in the model are related to each other and they can change when the economy goes from disequilibrium to equilibrium, so we let the prices endogenous in this section. In determining the prices of commodities and factors, we can link the production side (or supply side) to the consumption side (or demand side).

First of all, by using equations (5) and (6), we can obtain the price linkage between labour and capital.

$$P_L = \frac{\gamma_i}{(1-\gamma_i)}\frac{K_i}{L_i}P_K$$

Based on the zero economic profit assumption, the cost of producing X_i will determine the price of X_i, so,

$$P_i X_i = P_L L_i + P_K K_i = \frac{\gamma_i}{(1-\gamma_i)}\frac{P_K K_i}{L_i}L_i + P_K K_i = \frac{P_K K_i}{(1-\gamma_i)}$$

or

$$P_i = \frac{P_K K_i}{(1-\gamma_i)X_i}$$

Normalizing the price level of the economy to 1, we have

$$1 = \frac{\sum_{i=1}^{n+\Delta n} P_i X_i}{\sum_{i=1}^{n+\Delta n} X_i} = \frac{P_K \sum_{i=1}^{n+\Delta n} \frac{K_i}{1-\gamma_i}}{\sum_{i=1}^{n+\Delta n} X_i} \text{ or } P_K = \frac{\sum_{i=1}^{n+\Delta n} X_i}{\sum_{i=1}^{n+\Delta n} \frac{K_i}{1-\gamma_i}} = \frac{\sum_{i=1}^{n+\Delta n} (A_i + \Delta A_i) L_i^{\gamma_i} K_i^{1-\gamma_i}}{\sum_{i=1}^{n+\Delta n} \frac{K_i}{1-\gamma_i}},$$

or

$$P_i = \frac{K_i}{(1-\gamma_i)X_i}\frac{\sum_{i=1}^{n+\Delta n} X_i}{\sum_{i=1}^{n+\Delta n} \frac{K_i}{1-\gamma_i}} = (1-\gamma_i)^{-1}(A_i + \Delta A_i)^{-1} L_i^{-\gamma_i} K_i^{\gamma_i} \frac{\sum_{i=1}^{n+\Delta n} (A_i + \Delta A_i) L_i^{\gamma_i} K_i^{1-\gamma_i}}{\sum_{i=1}^{n+\Delta n} \frac{K_i}{1-\gamma_i}} \quad (20)$$

Based on this Pi, Ps can be easily calculated as $P_s = \sum_{i=1}^{n+\Delta n} P_i \delta_i$. Plugging Pi and Ps into equation (2), we can assess the impact of a change in inputs (K and L) on consumption (albeit a bit complicated).

The excess supply of commodities is the difference between commodities saved and commodities invested, i.e.

$$
\begin{aligned}
ES_i &= S_i - I_i = S\delta_i - I\beta_i \\
&= \delta_i (Y - \sum_{i=1}^{n+\Delta n} \theta P_i m_i + \sum_{i=1}^{n+\Delta n} \frac{P_i^2 \alpha_S}{2\alpha_i P_S}) / P_S - \frac{B}{1+r} \beta_i (\sum_{i=1}^{n+\Delta n} \theta m_i w_i - \sum_{i=1}^{n+\Delta n} (\theta m_i w_i - \frac{P_i \alpha_S w_i}{2\alpha_i P_S})) \\
&= \delta_i (Y - \sum_{i=1}^{n+\Delta n} \theta P_i m_i + \sum_{i=1}^{n+\Delta n} \frac{P_i^2 \alpha_S}{2\alpha_i P_S}) / P_S - \frac{B}{1+r} \beta_i \sum_{i=1}^{n+\Delta n} \frac{P_i \alpha_S w_i}{2\alpha_i P_S}
\end{aligned}
$$

The excess commodity supply will reduce the commodity supply in the next period, so the factors contributed to the excess commodity supply will not be employed in the next period. This causes unutilised labour and capital in the factor markets.

To avoid economic stagnation, it is necessary that there is no overall excess supply, i.e. $\sum ES_i \leq 0$. Summarizing all ES_i and using the fact that $\sum \delta_i = 1$ and $\sum \beta_i = 1$, we have:

$$
\sum_{i=1}^{n+\Delta n} ES_i = (Y - \sum_{i=1}^{n+\Delta n} \theta P_i m_i + \sum_{i=1}^{n+\Delta n} \frac{P_i^2 \alpha_S}{2\alpha_i P_S}) / P_S - \frac{B}{1+r} \sum_{i=1}^{n+\Delta n} \frac{P_i \alpha_S w_i}{2\alpha_i P_S} \leq 0
$$

$$
Y \leq \sum_{i=1}^{n+\Delta n} \theta P_i m_i + \frac{B}{1+r} \sum_{i=1}^{n+\Delta n} \frac{P_i \alpha_S w_i}{2\alpha_i} - \sum_{i=1}^{n+\Delta n} \frac{P_i^2 \alpha_S}{2\alpha_i P_S}
$$

This is the same results as equation (11) derived for the income ceiling from the demand-side perspective. However, the prices for commodities and for savings in the above equation are determined by the amount of capital and labour inputs through equation (20). Considering this, the ceiling on income is not fixed.

To be more accurate, the income derived above is measured by money so it is a nominal income. In real terms, the income in the equilibrium should be the goods consumed and invested, i.e.,

$$
Y_{real} = \sum_{i=1}^{n+\Delta n} \theta m_i + \frac{B}{1+r} \sum_{i=1}^{n+\Delta n} \frac{\alpha_S w_i}{2\alpha_i} - \sum_{i=1}^{n+\Delta n} \frac{P_i \alpha_S}{2\alpha_i P_S} \qquad (21)
$$

In this equation, the prices (and thus the labour and capital inputs) only affect the negative item, so the fixed ceiling still exists for real income.

Equation (21) can be used to demonstrate a case of government tax policy. Although there is no government in the model, the effect of a lump-sum tax can be demonstrated indirectly. If the government imposes this tax and distributes it to poor households, the distributional parameter θ in equation (21) increases with other things being equal, and this leads to an increase in real income. If the government use the tax to boost budget balance (i.e. tax revenue is unspent), this indicates an increased preference to save, so αS increase while αi decreases in equation (21). The consequence of this is an increase in the size of both positive investment effect (the second term at the right of equation) and the negative saving effect (the last term at the right of the equation) will increase. The enlarged negative saving effect is due to the decreased current consumption level due to government tax; and the enlarged positive investment effect is due to the increased gap between consumption ceiling and decreased current consumption. The over effect of this lump sum tax depends on the relative size of changes in both saving effect and investment effect. Generally speaking, the investment effect is smaller because the value of B / (1+r) is less than one, so the overall effect would be a decrease in real income.

5.3. Results from the recursive and intertemporal models: a dynamic perspective

The mechanism in the recursive model is essentially the same as that in the static model, so the equilibrium and disequilibrium in each period in the recursive model will be very similar to those in the static model. However, the investment function (7) plays a key role in pattern of economic growth. As the investment at time t (I_t) is proportional to the potential of consumption growth (i.e. the gap between current consumption and consumption ceiling), the change of potential of consumption growth leads to cyclic investment behaviour. When the gap between current consumption and consumption ceiling is large, investors perceive the high potential of increase in sales in the future and thus invest more in production. This leads to large increase in aggregate demand and pushes the economy into the boom phase. As the gap between current consumption and consumption ceiling gets smaller (assuming no new product is invented), an investment decrease is coupled by stagnancy or very slow growth of consumption. This leads to a decrease in aggregate demand and thus an economic recession.

During a recession, the firm cannot find a chance to increase sales because consumption level is close to consumption ceiling. Under this circumstance, firms have not choice but to invest in research and innovation, hoping to invent a new product. Once the innovation succeeds, the new product will lift the consumption ceiling and the gap between current consumption and consumption ceiling increase. As a result, both consumption and investment increase and the economy enter a recovery phase which is followed by an expansion phase. This cyclical growth will continue as long as there is no mechanism to stimulate product innovation.

From a different perspective, the intertemporal equilibrium model in the previous section can demonstrate the same point: a steady or balanced growth is not achievable in the long run if product innovation cannot keep up the pace of production growth.

We start with a task to find out the steady-state conditions as well as the optimal economic growth path. At the steady state, the growth of stock and capital becomes zero. Using equations (9) and (10), we have:

$$s' = af(k) - c - b(m - c) = 0$$

$$k' = b(m - c) - \delta k - nk = 0.$$

Combining the above two equations, we have:

$$k' = af(k) - c - \delta k - nk \quad (22)$$

Setting k' = 0, we have

$$c = af(k) - (\delta + n)k \quad (23)$$

Equation (23) defines a k’ = 0 curve.

The condition for optimal consumption is $\partial c / \partial k = 0$. This gives the same golden rule as in the Ramsey/Solow model:

$$af(k)' = (\delta + n) \quad (24)$$

At a steady state, the growth of consumption must also be zero. Setting equation (8) to zero, we have:

$$c' = -0.5e^{\theta t}[abf(k)' + (1 - b)(\delta + n)] + (\alpha_C \times m - \alpha_C \times c - 0.5\alpha_S)(\delta + n + \theta) = 0 \quad (25)$$

Equation (25) defines a c’ = 0 curve.

Combining equations (23) and (25), we can solve for the steady state E(c*, k*). However, this steady state will not be steady if the consumption ceiling (m) is fixed: due to the term $e^{\theta t}$, c’ will reduce over time. In other words, to maintain a steady state, m has to increase over time.

If the steady state is optimal, it must coincide with the optimal consumption point[3]. We apply the golden rule by substitute equations (23) and (24) into equation (25),

$$c' = -0.5e^{\theta t}(\delta + n) + (\alpha_C \times m - \alpha_C \times c - 0.5\,\alpha_S)(\delta + n + \theta) = 0 \quad (26)$$

This gives: $c = m - 0.5e^{\theta t}(\delta + n) / [\alpha_C(\delta + n + \theta)] - 0.5\alpha_S / \alpha_C$ (27)

The condition to achieve a steady consumption is:

$$dc / dt = dm / dt - 0.5\theta e^{\theta t} (\delta + n) / (\delta + n + \theta) = 0, \text{ or}$$

$$m = 0.5e^{\theta t} (\delta + n) / [\alpha_C(\delta + n + \theta)] + \text{constant}. \quad (28)$$

These results can be shown in Figure 3. Panel (a) shows the space view of the phase diagram. The k’ = 0 curve is given by setting k’ = af(k) – c – (δ + n)k = 0, i.e. equation (23). Based on equation (22), $\partial k' / \partial c = -1 < 0$, k’ and c move in opposite directions, i.e. as c increase, k’ decrease. This necessitates k’ > 0 within the k’ = 0 curve and k’ < 0 outside the k’ = 0 curve. Putting it differently, k will increase within the k’ = 0 curve and k will decreases outside the k’ = 0 curve. This movement is indicated by the solid arrows accompanied by letter ‘k’.

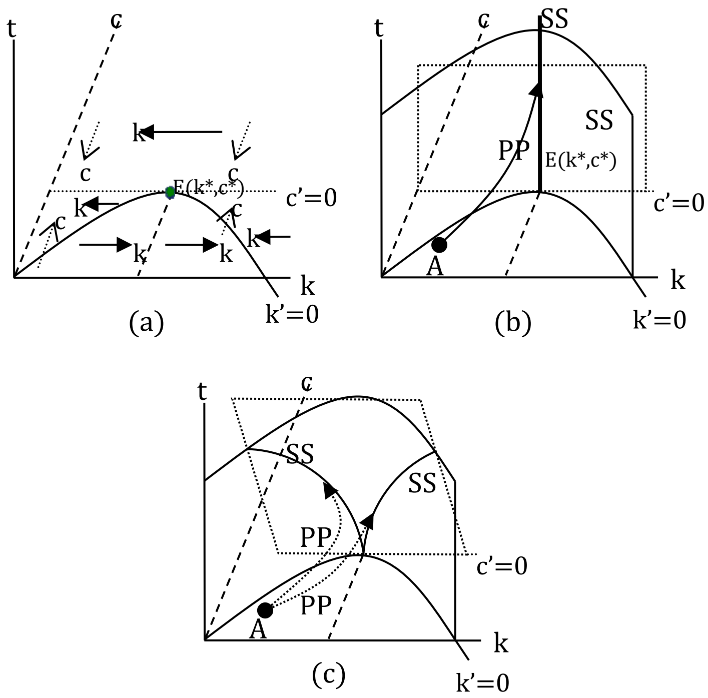
The steady state is at E(c*, k*), which is the intersection of the k’ = 0 curve and the c’ = 0 curve. For simplicity, we assume that the steady state is the optimal steady state and the c’ = 0 curve is c = c* (this assumption can be relaxed and the analysis is similar, but the graphs will be more complicated). According to equation (26), c’ and c move in opposite directions. As a result, c’ > 0 when c < c*, and c’ < 0 when c > c*. In other words, c will increase when c < c*, and c will decrease when c > c*. The movement of c is indicated by the dotted arrows accompanied by letter ‘c’.

The phase diagram for c and k indicates that the economy can converge to a steady state E(c*, k*) either (i) when c < c*, k < k*, and within the k’ = 0 curve, or (ii) when k > k* and outside of the k’ = 0 curve.

Panel (b) of Figure 3 shows the evolution of the economy over time. k' = 0 is now a curve space and c' = 0 is a vertical plane at c = c*. The tangent line of these two spaces SS shows the steady state of the economy. An economy at point A can reach the steady state SS through a path PP.

However, the $e^{\theta t}$ term in equation (28) shows that the existence of the steady state is conditional on the lifting of the consumption ceiling over time. If the consumption ceiling (m) is fixed, the term $e^{\theta t}$ in equation (26) necessitates that c will decrease over time. As a result, the c' = 0 plane will be inclined to the k axis and intersect with k' = 0 space on two curves SS1 and SS2. The economic equilibrium will evolve along either curve. The time path of economic growth is shown as the dotted arrow PP1 or PP2. Since the c and k on either SS1 or SS2 will change over time, there is no steady state – the consumption will decrease continuously.

Figure 3. *The phase diagram and economic growth path*



In short, the single-commodity intertemporal equilibrium model shows that, thanks to the fixed consumption ceiling, the economy will not reach a steady state – the consumption will keep falling in the long run. To reach a steady state, the consumption ceiling must keep increasing. Since the consumption of any commodity has a fixed ceiling according to the Axiom 1, the only way to increase the consumption ceiling for the economy is to increase the variety of commodities. That is, product innovation is the key to reach a steady state, or balanced growth.

5.4. The way to avoid a boom-bust cyclic growth

The analysis based on the static and dynamic models indicates that the only way to avoid an economic recession is to lift the income ceiling over time. These models also show that the key is to increase the number of commodity types, i.e. to invent new products. The links among consumption satiation, innovation and economic growth have been identified by previous researchers (e.g. Pasinetti, 1981; Andersen, 2001; Witt, 2001; Ruprecht, 2005), but no one has realised that product innovation is the key to avoid an economic recession and is the engine for economic growth. The current study shows that the repeated occurrence of economic recessions may indicate that product innovation is not able to lift the cap on income fast enough to avoid a recession. Why does innovation tend to lag behind? The obstacles to innovation activities are to blame.

One such obstacle is the high risk of investment in innovation. Innovation by definition involves the creation of something new. Inventors are continuously stepping into uncharted territory, so it is understandable that many successful innovations come only after numerous failed experiments. Although statistics on innovation failure/success are difficult to obtain, it is widely accepted that a high percentage of investment for research and development is not successful. There are two types of innovations: product innovation and process innovation. The goal of the former is to invent new products while the effort of the latter is to improve production efficiency. Compared with process innovations, product innovation has a much higher risk of failure because it normally involves much larger (or more radical) changes and there is much less information available to investors.

The other factor hindering innovation is imitation. Innovation requires hard and intelligent work, takes a long time, and requires a great deal of money. Imitating an innovation is, however, fairly easy. For example, software that takes several years and costs millions of dollars to develop, may take only a few minutes to copy. Other imitations may be harder and cost more (e.g. re-engineering a medicine) but, in considering the numerous failures before a successful invention and the correct directions an imitator learned from the successful inventor, the imitation cost will be much lower than the cost of invention, As a result, the externality of inventions should be large.

Product innovation is much more vulnerable to imitation than process innovation. Because process innovations are applied to production procedures or machinery, imitating these innovations requires knowledge about the production environment. Imitating a new product does not, however, require this knowledge. The vulnerability of product innovation to imitation means that the externality of product innovation is enormous.

Due to the distinct possibility of innovation failure and the low chance of getting a good return because of imitation, risk-averse investors are reluctant to invest their money in innovations, especially in product innovation. Rather, they prefer to invest in production that has a relatively certain investment return. Therefore, innovation investments, or R&D funds, become scarce, and this leads to scarcity of product innovation activity.

The scarcity of product innovation leads to a cap on income (Y in equation 11) and thus a cap on economic growth. When the household tries to break the cap on income by supplying more labour and capital to the firm, the firm will produce more commodities than can be cleared by the market. This will cause a recession. During an economic recession or stagnation, it is not profitable to invest in production because the goods are unsellable, so investors are forced to invest in product innovation in the hope of obtaining a profitable new product. As such, 'Δn' increases, the cap on income is lifted, and the economy starts to recover and expand. This recession-recovery-expansion cycle will continue and result in cyclic economic growth.

It is worth mentioning that the importance of innovation to an economy is well recognised by most people. Most influential study is Kondratiev (1922) and Schumpeter (1939), which link innovations to business cycles. There are also numerous studies on product and process innovations. Notable studies include Lancaster (1966), Romer (1990), Grossman and Helpman (1991), Aghion and Howitt (1992), Jones (1995), Klette and Kortum (2004), Acemoglu et al. (2013), and Georges (2015). In these studies, innovations are endogenously determined by human capital or R & D funds, so innovation scarcity is not concern of these studies. Although Schumpeter (1939) attributed economic recession to the lack of innovation, no one so far has identified or discussed the cause of innovation scarcity and no one realized that the scarcity of product innovation is the root cause of economic recessions.

Is there any way to stimulate investment in product innovation so as to avoid economic recessions? Although no one can change the high-risk nature of product innovation, one can balance this high risk with high return through forbidding imitation. An effort of this kind is seen in the patent laws, but the limitations in current patent laws lead to the failure to protect inventors fully. A full discussion of the limitations of patent laws is to be covered by another paper, here we discuss them only very briefly. On one hand, current patent laws impose limited durations and a compulsory license rule on patent rights, aiming at moderating the monopoly power of the patentee and at forcing the patentee to implement patented technology. These clauses cause considerable stress for inventors and discourage innovation. On the other hand, the patent laws allow granting exclusive patent licenses, which transfer monopoly power of patentees to licensees so that the patent monopoly power magnifies in the economy. This aggravates the problem of abusing patent monopoly power.

A thorough revision of patent laws can further encourage innovation activity while overcoming the problem of patent monopoly power abuse in a positive way. For example, abolishing the time constraint put on patentees, forbidding both the assignment of patent rights and exclusive patent licensing (to avoid the magnification of patent monopoly power), and implementing a non-exclusive patent licensing system. Under this system, anyone could use the innovation by applying for a license from the inventor. With the property right of the patent licensing established and clearly defined (infinite duration of patent right) in the new system, a patent market may come into reality and it will automatically channel funds into innovation activities.

6. Conclusions

Based on observations and reasoning in the real world, this paper proposed three axioms. (1) There is a consumption satiation point for each commodity per household. (2) Savings can generate utility for households directly and immediately. (3) Investment demand is critically determined by profitability and, at the macroeconomic level, the latter is indicated by the growth of household consumption.

In including these axioms in a multi-commodity model, the paper uncovers the ultimate cause of economic recessions. An economic recession can be attributed to the stagnation of household consumption due to the existence of consumption satiation points and due to the scarcity of product innovation. Since the profit for investment eventually comes from the sales to households, consumption stagnation leads to investment decline. The stagnation in household consumption and the decrease in investment demand work together to drive the economy into a recession.

The paper also explains the cyclic economic growth and the vital role of product innovation – the product innovation cycle underpins the investment cycle and thus a cyclic economic growth pattern. This signals that product innovation scarcity is a major obstacle to long-run economic development. The discussion on the reasons for innovation scarcity and on the limitations in current patent laws points to the necessity to revise these laws. If more effective and efficient patent laws would lead to the formation of a functional patent market, the synergy of the capital market and the patent market can maintain an appropriate balance between investment in product innovation and investment in production, and thus the pace of product innovation is balanced with the pace of production. As a result, economic recessions can be avoided and faster and smoother rates of economic growth can be achieved.

Notes

(1) It is arguable that consumption can include commodities purchased for future use or for others' use (e.g. gift, donation or bequeathing). This type of consumption can be treated either as other people's consumption if the commodity is used up, or as other people's savings if the commodity is unused by other people, which will be discussed next.

(2) For simplicity, lending and borrowing are not considered in this paper. Lending and borrowing can delay the problems caused by consumption ceiling but cannot change the nature of the consumption ceiling because debts are required to be paid off eventually. Explicitly including lending and borrowing will not change the results, but will complicate the model.

(3) A steady state can happen at any capital level, i.e. multiple steady states. In this case, we can obtain at each capital level a consumption level at which c' = 0. The analysis for each steady state is similar. The assumption of an optimal steady state simplifies the analysis.

References

Acemoglu, D., Akcigit, U., Bloom, N. and Kerr, W., 2013. Innovation, Reallocation and Growth, NBER WP 18933.

Aghion, P. and Howitt, P., 1992. A Model of Growth Through Creative Destruction, *Econometrica,* 60(2), pp. 323-351.

Aghion, P. and Howitt, P., 2006. Joseph Schumpeter Lecture Appropriate Growth Policy: A Unifying Framework, *Journal of the European Economic Association,* 4(2-3), pp. 269-314.

Akerlof, G. and Yellen, J. 1985. Can Small Deviations from Rationality Make Significant Differences to Economic Equilibria? *American Economic Review,* 75(4), pp. 708-720.

Andersen, E., 2001. Satiation in an Evolutionary Model of Structural Economic Dynamics, *Journal of Evolutionary Economics,* 11(1), pp. 143-164.

Aoki, M. and Yoshikawa, H., 2002. Demand Saturation-creation and Economic Growth, *Journal of Economic Behaviour and Organization,* 48(2), pp. 127-154.

Cass, D., 1965. Optimum Growth in an Aggregative Model of Capital Accumulation, *Review of Economic Studies*, 32, pp. 233-240.

Chai, A. and Moneta, A., 2014. Escaping Satiation Dynamics: Some Evidence from British Household Data, *Journal of Economics and Statistics,* 234 (2/3), pp. 299-327.

Davidson, P., 1984. Reviving Keynes's Revolution, *Journal of Post Keynesian Economics,* Fall.

Davidson, P., 1991. Is Probability Theory Relevant for Uncertainty? A Post Keynesian Perspective, *Journal of economic perspectives,* Winter.

European Commission, 2013. Unemployment Statistics, <http://epp.eurostat.ec.europa.eu/statistics_explained/index.php?title=File:Unemployment_rate,_2001-2012_%28%25%29.png&file timestamp=20130627102805>

Fisher, I., 1930. *The theory of interest,* New York: The Macmillan Co.

Fisk, E., 1962. Planning in a Primitive Economy: Special Problems of Papua New Guinea, *Economic Record,* 38, pp. 462-478.

Friedman, M., 1957. *A Theory of the Consumption Function,* Princeton University Press.

Georges, C., 2015. Product Innovation and Macroeconomic Dynamics, Manuscript. Hamilton College. <http://academics.hamilton.edu/economics/cgeorges/product-innovation-and-macro-dyn.pdf>

Grossman, G. and Helpman, E., 1991. *Innovation and Growth in the Global Economy.* Cambridge, MA: The MIT Press

Howe, H., 1975. Development of the Extended Linear Expenditure System from Simple Saving Assumptions, *European Economic Review,* 6, pp. 305-310.

Hutcheson, F., 1750. *Reflections upon laughter, and remarks upon the Fable of the bees,* Printed by R. Urie, for D. Baxter

Jones, C., 1995. R & D-Based Models of Economic Growth, *Journal of Political Economy,* 103(4), pp. 759-784.

Keynes, J.M., 1936. *The General Theory of Employment, Interest, and Money.* London: MacMillan and Co.

Kimball, M., 1990. Precautionary Saving in the Small and in the Large, *Econometrica,* 58(1), pp. 53-73.

Klette, J. and Kortum, S., 2004. Innovating firms and Aggregate Innovation, *Journal of Political Economy.* 112(5), pp. 986-1018.

Kondratiev, N.D., 1922. The World Economy and its Conjunctures during and After the War (in Russian), Moscow.

Koopmans, T., 1965. On the Concept of Optimal Economic Growth, in *The Econometric Approach to Development Planning,* North Holland, Amsterdam.

Kydland, F.E. and Prescott, E.C., 1982. Time to Build and Aggregate Fluctuations, *Econometrica,* 51, pp. 1345-1370.

Lancaster, K., 1966. Change and Innovation in the Technology of Consumption, *American Economic Review,* 56, pp. 14-23.

Lluch, C., 1973. The Extended Linear Expenditure System, *European Economic Review*, 4, pp. 21-32.

Lucas, R.E., 1975. An Equilibrium Model of the Business Cycle, *Journal of Political Economy,* December.

Mandeville, B., 1723. *The fable of the bees, or, Private vices, public benefits.* 2nd ed., London: Printed for Edmund Parker.

Mankiw, N.G., 1985. Small Menu Costs and Large Business Cycles: A Macroeconomic Model of Monopoly, *The Quarterly Journal of Economics,* Vol. 100, No. 2 (May), pp. 529-537.

Mankiw, N.G., 1989. Real Business Cycles: a New Keynesian Perspective, *Journal of Economic Perspectives,* Summer.

Modigliani, F., 1986. Life cycle, Individual Thrift, and the Wealth of Nations, *American Economic Review,* 76, pp. 297-313.

Modigliani, F. and Brumberg, R. 1954. Utility Analysis and the Consumption Function: An Interpretation of Cross-section Data, in K. Kurihara ed., *Post-Keynesian Economics,* Rutgers University Press, New Brunswick.

Pasinetti, L., 1981. *Structural Change and Economic Growth: a Theoretical Essay on the Dynamics of the Wealth of Nations,* Cambridge University Press, Cambridge.

Pigou, A.C., 1943. The Classical Stationary State, *Economic Journal*, 53, pp. 343-351.

Plosser, C.I., 1989. Understanding Real Business Cycles, *Journal of Economic Perspectives,* Summer.

Prescott, E.C., 1986. Theory Ahead of Business Cycle Measurement, Federal Reserve Bank of Minneapolis Quarterly Review, Fall.

Ramsey F., 1928. A Mathematical Theory of Saving, *Economic Journal*, 38, pp. 543-559.

Romer, D., 1993. The New Keynesian Synthesis, *Journal of Economic Perspectives,* Winter.

Romer, D., 2013. *Advance Macroeconomics,* 4^{th} edition, The McGraw Hill companies.

Romer, P., 1990. Endogenous Technological Change, *Journal of Political Economy,* 98(5), pp. S71-S102.

Ruprecht, W., 2005. The Historical Development of the Consumption of Sweeteners: a Learning Approach, *Journal of Evolutionary Economics,* 15(3), pp. 247-272.

Saviotti, P., 2001. Variety, Growth and Demand, *Journal of Evolutionary Economics,* 11(1), pp. 119-142.

Saviotti, P. and Pyka, A., 2013. The Co-evolution of Innovation, Demand and Growth, *Economics of Innovation and New Technology,* 22(5), pp. 461-482.

Schumpeter, J., 1939. Business Cycles: *A Theoretical, Historical, and Statistical Analysis of the Capitalist Process,* New. York and London, McGraw-Hill Book Company Inc.

Smith, A., 1776[1904]. *An Inquiry Into the nature and Causes of the Wealth of Nations,* Edwin Cannan ed., London: Methuen.

Sorensen, P. and Whitta-Jacobsen, H., 2010. *Introducing Advanced Macroeconomics* (2nd ed.), McGraw-Hill, London.

Stent, W. and Webb, L., 1975. Subsistence Affluence and Market Economy in Papua New Guinea, *Economic Record,* 51, pp. 522-538.

Tobin, J., 1969. A General Equilibrium Approach to Monetary Theory, *Journal of Money Credit and Banking*.

Turkington, D., 2007. *Mathematical Tools for Economics,* Blackwell Publishing.

Witt, U., 2001. Escaping Satiation: the Demand Side of Economic Growth, Springer.