Say's Law and the Business Cycles
Xianming Meng

School of Economics, University of New England, Australia
E-mail: mxmbox@yahoo.com

The relationship between Say's Law and the business cycle has been vigorously debated but no generally accepted agreement has yet been reached. Although this is a disputed topic, revisiting it anew could further develop economic theory, and enhance the further growth of the global economy. By scrutinising Say's Law, and its implicit business cycle theories, this paper claims that general gluts are possible, and that the essence of the business cycle is in innovation scarcity. Thus in order to reduce the effects of the business cycle and spur economic growth, a thorough revision of patent laws is necessary.

Key Words: Say's law; Innovation; Business cycle; Patent.
JEL Classification Numbers: E19, E32.

1. INTRODUCTION

Say's Law (or the Law of Markets) and the business cycle have been heavily debated throughout economic history, but they seem of importance only to historians of economics today. This is because economists believe that, although no consensus about these topics has yet been reached, they fully understand them. In addition, Say's Law is no longer among the conceptual tools habitually employed by economists. Finally, business cycles have become less frequent and less severe. Nevertheless, this paper will argue that, in spite of fewer large fluctuations, the post-industrial economy may be in low-growth trap, and that Say's Law, albeit with some emasculation, still contains some clues concerning the business cycle.

The paper contains seven sections. The next section is titled, the original incarnation of Say's Law, which will trace the origin of Say's Law and expose the flaws in its reasoning. Section 3, modern interpretations of Say's Law, will briefly introduce and comment on some modern representative restatements of Say's Law. In section 4, the business cycle theories implied by Say's Law will be analysed. Section 5, the essence of the business cy-
section, will consider the causes of the business cycle. Section 6 will discuss a method of alleviating the negative effects of the business cycle and increasing economic growth. In final section, conclusions will be drawn based on the said analysis.

2. THE ORIGINAL INCARNATION OF SAY’S LAW

The controversy around Say’s Law has continued for about two centuries. The most unusual debate concerns the origin of Say’s Law. This debate is of the least economic value; but by tracing the origin of Say’s Law, we can have a better understanding of the true meaning of Say’s Law.

A. The origin of the term “Say’s Law”

Say’s Law became well-known with the publishing of Keynes’s famous book General Theory, in which it was comprehensively refuted. However, there was no knowledge about how and why it was named after the French economist Jean-Baptiste Say until Kates (1995) discovered that it was Fred Taylor (1925) who first coined the term Say’s Law. In Chapter XV of his book Principles of Economics, Taylor (1925, p.201) wrote:

I shall therefore put the proposition we have discussed in the form of a principle. This principle, I have taken the liberty to designate Say’s Law; because, though recognized by many earlier writers, it was particularly well brought out in the presentation of Say (1803).

However, in the first edition of Treatise (1803), Say did not express such a law explicitly. When Say’s Law is considered reference is usually made to the second edition Treatise (1814) or the English translation of the fourth edition Treatise (1821), in which Say (1821, Vol. I, Book I, p.167) provided his widely cited and generally accepted expression of Say’s Law:

A product is no sooner created, than it, from that instant, affords a market for other products to the full extent of its own value ... the mere circumstance of the creation of one product immediately opens a vent for other products.

Subsequent to this expression, despite considerable amplifications, implications and applications of Say’s Law, here we can safely simplify the content of Say’s Law as “the act of supply creates an equivalent demand for other products”.

B. The originator of Say’s Law

Although some economists (e.g. Spengler, 1945; Winch, 1966; and Sowell, 1972) agree that Say was the originator of Say’s Law, not all economists support this idea. Some (e.g. Thweatt, 1979; Kates, 1997; and Baumol, 1999) argue that James Mill was the true author of the Law of Markets. Some go future to maintain that Adam Smith was the father of the law, and others even suggest that the origin of Say’s Law can be traced further back. It is instructive to consider what Smith and Mill said in their books.
In the Wealth of Nations, Adam Smith (1776, p268, 321) said:

In all countries where there is tolerable security, every man of common understanding will endeavor to employ whatever stock he can command, in producing either present enjoyment or future profit ... A man must be perfectly crazy who, where there is tolerable security, does not employ all the stock which he commands, whether it be his own or borrowed of other people, in some one or other of those ... ways (p268).

Later, he adds:

That portion of his revenue which a rich man annually spends, is in most cases consumed by idle guests, and menial servants, who leave nothing behind them in return for their consumption. 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 (p321).

Smith’s expression here can be summarised as the famous theorem that savings are spent as quickly as consumption. This implies that since all income (consumption and savings) is spent, total demand equals total supply. However, it does not imply that supply creates demand or that supply determines demand, which is the core of Say’s Law. So we should by no means attribute Say’s Law to Smith.

In Commerce Defended James Mill (1808, p.81) wrote:

No proposition however in political economy seems to be more certain than this which I am going to announce, how paradoxical soever it may at first sight appear; and if it be true, none undoubtedly can be deemed of more importance. The production of commodities creates, and is the one and universal cause which creates, a market for the commodities produced. Let us but consider what is meant by a market. Is anything else understood by it than that something is ready to be exchanged for the commodity which we would dispose of? When goods are carried to market what is wanted is somebody to buy. But to buy, one must have wherewithal to pay. It is obviously therefore the collective means of payment which exist in the whole nation that constitute the entire market of the nation. But wherein consist the collective means of payment of the whole nation? Do they not consist in its annual produce, in the annual revenue of the general mass of its inhabitants? But if a nation’s power of purchasing is exactly measured by its annual produce, as it undoubtedly is; the more you increase the annual produce, the more by that very act you extend the national market, the power of purchasing and the actual purchases of the nation ... .

Mill’s words explicitly express the concept that production creates a market. His explanation of the market here corresponds with the generally accepted idea of demand. Thus his expression that production creates demand does predate Say’s similar proposition in the second edition of Treatise.

However, before Mill is credited with the origin, Say’s first edition of Treatise should be considered. Almost nothing can be found about the Law of Markets in the famous Chapter 22, “Of Markets”, as James (1808), Spengler (1945) and Baumol (1977) have claimed. However there are per-
tinent discussions of the Law of Markets in other chapters. For instance, in Chapter 5, Say (1803, Vol. II, Book 4, p.175) wrote:

In order to consume it is necessary to purchase; now, one can make purchases only with what one has produced. Is the quantity of outputs demanded consequently determined by the quantity of products created? Without any doubt. Everyone can, at his pleasure, consume what he has produced; or else he can buy another product with his own. The demand for products in general is therefore always equal to the sum of the products available. (italics added)

Here Say claims that demand is determined by production so total demand is always equal to total supply. Although this expression is not as explicit and systematic as Mill’s, it is years earlier. We should also acknowledge another fact: that James Mill had already read Say’s treatise when he wrote Commerce Defended. Taking account of all these data, Fred Taylor was correct to claim that Say was the father of Say’s Law.

C. Fatal flaws in Say’s Law

Although Say’s Law is based on the simple and important fact that the supply creates the income which is necessary for demand, it is not entirely correct due to some fatal flaws in its reasoning.

(a) Mistaking the demand for the purchasing power

From Say’s widely cited expression of Say’s Law Say’s reasoning be can followed: because supply creates income and thus purchasing power, the action of supply creates an equivalent demand for other products. Apparently, here Say did not distinguish demand from purchasing power. To avoid the distinction between demand and purchasing power, he used the words “market” or “vent” instead of “demand”. The same approach occurred in Mill’s expression of Say’s Law. When defining the meaning of the market, Mill treated “the national market, the power of purchasing and the actual purchases of the nation” as the same.

A failure to differentiate demand from purchasing power is the fundamental flaw in Say’s Law. If demand means purchasing power, or the quantity demanded always equals purchasing power, the reasoning in Say’s law is absolutely correct. But if not, Say’s law becomes untenable. That is why even Mill (1808, p.81) felt the propositions of Say’s Law may at first sight appear paradoxical.

Interestingly, at the very beginning, Ricardo noticed the differences in meanings of demand. As early as 1811, in a letter replying to Mill, Ricardo (1811, Vol. 6, p.56) wrote:

You observe that the demand for corn is unlimited. It is clear that you attach a different meaning to the word demand to what I do. I should not call the mere desire of possessing a thing a demand for it ... By demand I should understand a desire to possess with the power of purchasing'.

Mill replied: “I follow Dr Smith’s rule, which is to call it effectual demand, as often as it means the will to purchase combined with the power.”
SAY’S LAW AND THE BUSINESS CYCLES

(Ricardo, 1811, Vol. 6, p.58) Later, Mill (1844, p.230-231) developed the idea that demand and its equivalent (purchasing power) are convertible terms:

Two things are necessary to constitute a demand. These are, first, a wish for the commodity; second, an equivalent to give for it. A demand means the will to purchase, and the means of purchasing. The extent of his demand is measured by the extent of his equivalent. The demand and the equivalent are convertible terms, and the one may be substituted for the other. The equivalent may be called the demand, and the demand the equivalent.

The distinction between the will and the means to purchase is apparent, but among economists it is argued whether the will to purchase falls short of the power to purchase. In their correspondence, Malthus declared to Ricardo: “I by no means think that the power to purchase necessarily involves a proportionate will to purchase.” Ricardo replied that, “We agree too that effectual demand consists of two elements, the power and the will to purchase, but I think the will is very seldom wanting where the power exists” (Ricardo, 1814, Vol. 6, p.133, italics added). The argument continued, as can be seen here, where McCulloch (1864, p.146) joins the argument against Malthus:

Malthus has justly stated that the demand for a commodity depends “on the will combined with the power to purchase it” that is on the power to furnish an equivalent for it. But who ever heard of a want of will to purchase? If it alone could procure necessaries and luxuries, every beggar would be rich as Croesus, and the market would constantly be understocked. The power to purchase is the real desideratum. It is the inability to furnish equivalents for the products necessary to supply our wants that “makes calamity of so long a life”.

Is the will to purchase always as unlimited as McCulloch suggested? To answer this question it is necessary to explore the meaning of “the will to purchase”. This term refers to individuals’ desire to consume commodities or services. Generally speaking, the human desire to consume is unlimited. Human beings are unsatisfied creatures — they want a better life no matter how much their life has already been improved. It is this unsatisfied nature that drives human progress. However, the human desire for any kind of commodity or service is definitely limited. For example, people cannot eat unlimited food every day and cannot have their homes cleaned and decorated simultaneously.

There seems some contradiction here: if the desire for each commodity or service is limited, the desire for all of them should be limited as well. This gap in the reasoning can be explained by the variety of commodities and services. If the variety of them is limited, the total desire for them is also limited (despite the profligacy of some people). On the other hand, if they are unlimited in number, the total desire to consume is unlimited as well. In the long run, with new commodities or services surfacing, the variety of goods will keep increasing, and so human demand for them will
be unlimited. But in the short term, the variety of goods is limited, so the will to purchase them is not unlimited, and thus may fall short of the power to purchase.

(b) Equating saving with investment

It is commonsense that, between individuals, the will and the power to purchase vary. For the poor, the will to purchase is larger than their power to do so; for the rich, it is the opposite situation. If the nation as a whole is considered, does the total will to purchase equal the total purchasing power? The key to the answer lies in the national savings.

National income (national purchasing power) can be divided into two parts: consumption and savings. National consumption is, by definition, spent income. If national savings are not spent, the total will to purchase will be less than the national income, and general gluts, or under-consumption, may occur. Reacting to this concern, Adam Smith (1776 [1904], Vol. I, p.320-321) put forward the famous “saving is spending” theorem:

What is annually saved is as regularly consumed as what is annually spent, and nearly in the same time too. By saving a part of [the rich man’s revenue], as that part is for the sake of profit immediately employed as a capital either by himself or some other person, the food, clothing, and lodging, which may be purchased with it, are necessarily reserved for the latter. The consumption is the same, but the consumers are different.

In figure 1, vertical axis $i$ refers to the interest rate, and horizontal axis $q$ stands for the quantity invested or saved. $I$ is the investment curve which shows total investment and interest rate combinations, and $S$ is the saving curve, showing level of saving at different interest rates. Initially, investment and saving are at equilibrium when $i = i_0$, namely, $S = I = q_0$. Assume that for some reason (for example national income increases) national saving increases; the saving curve shifts right to $S'$. At interest rate $i_0$ there exists unexploited saving $q' - q_0$. However, since saving is greater than investment, the interest rate will fall. Eventually, at a lower interest rate $i_1$, saving and investment will be back to equilibrium at $q_1$.

It seems that interest rates can balance saving and investment automatically. However, the problem is that investment is not only determined by the interest rate. If an investment does not make a profit or even incurs a loss, it is commonsense not to invest even if the interest rate is zero. This phenomenon deeply concerned Say.
In his Letters to Mr Malthus, Say (1820 [1821], p.36n) calls attention to the failure of demand.

what happened to us in 1813 ... when interest of money fell so low, for want of good opportunities of employing it and by what is happening to us at this moment in which the capitals sleep at the bottom of the coffers of the capitalists.

That investment may not always be profitable causes the main problem in the “saving is spending” theorem. As Baumol (1977, p.159) suggested, Smith should ‘have considered a period of bad business prospects or great uncertainty to involve intolerable security’. If we trace the cause of profitability of investment, we can find that, although investment (called “reproductive consumption” by Say), was thought by classical economists (including Smith, Mill, and Say) to be more important than final consumption, it is affected by the latter (called “unproductive consumption” by Mill and Say).

The purpose of investment is to earn profit by producing a final good or service. To achieve this end, the final good or service must find a market. If final consumption is weak, the final good will not be sold, and the investment will not be profitable or may even cause a loss. In this case, because an investment opportunity is not available, saving will not be employed as capital. In short, since the demand for capital goods depends on the demand for final goods, investment is dependant on final consumption.

(c) Supply according to demand
If it is presumed that demand equals purchasing power, it can also assumed that an individual works according to his demand (by this assumption an individual’s income from work will exactly satisfy his demand). James Mill (1844, p.228) makes a good argument about this.

When a man produces a greater quantity of any commodity than he desires for himself, it can only be on one account; namely that he desires some other commodity which he can obtain in exchange for the surplus of what he himself has produced. It seems hardly necessary to offer any thing in support of so necessary a proposition; it would be inconsistent with the known laws of human nature to suppose, that a man would take the trouble to produce any thing without desiring to have any thing.

This assumption sounds convincing - if an individual’s consumption is low, it is meaningless to work too hard; instead, he can have more leisure time. Nevertheless, there are some problems with this assumption. First, if an individual produces according to his anticipated consumption, all his needs will probably be unfulfilled. Since some products he wants to consume are produced by other individuals, his consumption depends not only on the quantity of goods he produces himself, but also on the quantity of goods produced by other individuals. If the good produced by another individual becomes very dear because it becomes popular, one will find that he has produced too few products to exchange for enough goods he wants. When he realises that his plan will not be fulfilled, the best thing to do is to produce more or simply give up on this strategy.

Second, this strategy is unpractical for an economy in that it may lead to shrinking production. If one’s demand is low, one will produce less. The less one produces, the more he can charge for his products so as to benefit from an exchange. If everyone in the economy has this knowledge, everyone will try to produce less, and the economy will shrink.

Third, some desires to consume vary unexpectedly, so they cannot be put into a plan. For instance, an individual may not plan to eat anything, but the smell from a food stand is so good that he makes a purchase. Even planned consumption may change considerably. When it comes to the highly elastic luxuries, spending may change wildly. A consumer may plan to buy a $14 000 car, but because of the influence of the car dealer, he actually spends $28 000. In short, people’s demands keep changing (that is why commercials can be influential), so at least some consumption cannot be anticipated.

Last, but most importantly, people may harbour some seemingly impossible desires in their mind, which later become attainable. For instance, traveling to Mars cannot be fulfilled today due to technology constraints. But if the constraints were removed, people could partake of the hitherto unprovided service. Since new products and possibilities will become available, people cannot have an exact plan for their future consumption.
To sum up, supply according to demand is not practical in an economy. On the contrary, what is evident is that firms produce according to their production capability (as long as they can sell their products) so as to maximise profit. Individuals, too, work in order to maximise income. By doing so, individuals ensure they satisfy current demands through exchange and can fulfill any potential demands if a new product or service they want becomes available.

3. MODERN INTERPRETATIONS OF SAY’S LAW

Since Say did not codify Say’s Law in his book, the content of Say’s Law became controversial. For instance, Baumol (1977) thought there were at least eight laws in Say’s statement. Thweatt (1979) concluded that four essential aspects exist in a full and mature statement of Say’s Law. In the long process of theoretical refinement, Say’s Law has both lost and acquired meanings. This section will introduce some representative interpretations of Say’s Law.

A. Keynes’s interpretation

Keynes’s interpretation, simplified as “supply creates its own supply”, is the most famous, widely cited and criticised one. In his famous book General Theory, Keynes (1936, p. 18-19) wrote:

From the time of Say and Ricardo the classical economists have taught that supply creates its own demand — meaning by this in some significant, but not clearly defined, sense that the whole of the costs of production must necessarily be spent in the aggregate, directly or indirectly, on purchasing the product ...

As a corollary of the same doctrine, it has been supposed that any individual act of abstaining from consumption necessarily leads to, and amounts to the same thing as, causing the labour and commodities thus released from supplying consumption to be invested in the production of capital wealth ...

According to Keynes’s interpretation of Say’s Law, any supply of commodities will create demand for themselves, so no commodity can be over-supplied. As a result, an economic system will always operate at full capacity, and economic recession and involuntary unemployment are totally impossible. As Keynes (1936, p.26) puts it:

Thus Say’s law, that the aggregate demand price of output as a whole is equal to its aggregate supply price for all volumes of output, is equivalent to the proposition that there is no obstacle to full employment.

Although Keynes’s interpretation is thought as being “the well known naive rendition” (Jonsson, 1995) of Say’s Law, there is little discrepancy with what Say says. If “the mere circumstance of the creation of one product immediately opens a vent for other products” (Say, 1821, Vol. I, Book I, p.167), then the reverse should also be true: the creation of other products will open a vent for it. Consequently, the creation of one product
itself opens its own vent, or, supply creates its own demand. In the case of a three-product economy, if supply of product A creates a demand for product B and supply of product B creates a demand for product C, the supply of product C will in turn create a demand for product A. It could be argued that Say’s expression allows a partial glut. For example, the supply of product A creates a demand for product B when A itself is oversupplied. If it is true, however, because of partial glut of A, the demand for product B will not reach its full value, which is contradictory with Say’s expression: “A product is no sooner created, than it, from that instant, affords a market for other products to the full extent of its own value” (1821, Vol. I, Book I, p.167). All in all, Keynes’s interpretation of Say’s Law is not a “straw man”. Rather, it is a logical deduction from Say’s expression.

B. Lange’s restatement

In Lange’s often cited and criticised paper “Say’s Law: a Restatement and Criticism” (1942), the opening words are: “Say’s law is the proposition that there can be no excess of total supply of commodities (general oversupply) because the total supply of all commodities is identically equal to the total demand for all commodities”. Lange proceeds to demonstrate this mathematically.

In a closed economy with $n$ commodities, let $P_i$, $D_i$, and $S_i$ be the price, demand function and supply of $i$th commodity respectively. Let the $n$th commodity be money and $p_n = 1$. By definition,

\[ \sum_{i=1}^{n-1} P_i D_i \equiv S_n \]  
(total demand for non-money commodities equals total supply of money)

\[ \sum_{i=1}^{n-1} P_i S_i \equiv D_n \]  
(total supply of non-money commodities equals total demand for money)

so total demand for all commodities including money is:

\[ \sum_{i=1}^{n} P_i D_i \equiv \sum_{i=1}^{n-1} P_i D_i + p_n D_n \equiv S_n + D_n \]

and total supply of all commodities including money is:

\[ \sum_{i=1}^{n} P_i S_i \equiv \sum_{i=1}^{n-1} P_i S_i + p_n S_n \equiv D_n + S_n \]

therefore, \[ \sum_{i=1}^{n} P_i D_i \equiv \sum_{i=1}^{n} P_i S_i \]  
(total demand for and supply of all commodities including money are identically equal)

The equation is actually an expression of Walras’s Law. However, Say’s Law claims that total demand for and supply of non-monetary commodities
are identically equal, namely:

\[ \sum_{i=1}^{n-1} P_i D_i \equiv \sum_{i=1}^{n-1} P_i S_i, \]  

which means \( D_n \equiv S_n, \) or \( \Delta M \equiv D_n - S_n \equiv 0. \)

Thus, according to Lange (1942, p.153), “Say’s Law implies a peculiar nature of the demand for money, namely, that the individuals in our system, taken together, are always satisfied with the existing amount of money and never wish to hold either more or less”.

This conclusion would not surprise Say if he had survived until modern times, because he thought money was nothing but an intermediary with which to facilitate exchange. In Chapter 22 of his Treatise (1803, Vol. I, Book 1, p.154), Say wrote, “Money performs no more than the role of a conduit in this double exchange. When the exchanges have been completed it will be found that one has paid for products with products.” Although Lange made a mistake about whether Walras’s Law or Say’s Law came first, his interpretation is valuable for it makes explicit the contradiction between Say’s Law and modern monetary theory. Lange (1940, p.167) claims: “Say’s Law precludes any monetary theory. The theory of money must, therefore, start with a rejection of Say’s Law”.

C. Say’s equality suggested by Becker and Baumol

In response to Lange’s challenge, Becker and Baumol in 1952 published the paper: The Classical Monetary Theory: The Outcome of the Discussion. In this paper they distinguished Walras’s Law, Say’s identity, and Say’s equality.

According to them, Walras’s law shows that the total value of goods (including money flow) demanded equals the total value of goods (including money flow) supplied. Say’s identity shows, at any set of prices, that the total monetary value of the demand for commodities (excluding money) will be equal to the total monetary value of the quantity of all commodities supplied (excluding money), assuming that individuals use the money they receive to demand other commodities immediately. Say’s equality refers to the proposition that, for every relative price structure, there exists a price level at which the money market will be in equilibrium. In a later paper Baumol (1977, p.159) reinforced the conclusion that Say’s Law means Say’s equality: “Thus the eighth (and for our purposes the last) of Say’s eight propositions is Say’s Law itself. Apparently this takes the form of a type of Say’s equality, i.e. supply and demand are always equated by a rapid and powerful equilibration mechanism”.

As a result of Becker and Baumol’s paper, it seemed that finally “classical economists have been absolved of being the fools that Keynes made them out to be” (Kates, 1997, p.195). Yet their argument is not flawless, for the
equalizing mechanism which Becker and Baumol thought was rapid and powerful is problematic. We will return to it later.

**D. Say’s Principle relabeled by Clower and Leionhufrud**

In Say’s Principle, What It Means and Doesn’t Mean, Clower and Leionhufrud (1973, p.80) relabeled Say’s Law as Say’s Principle (SP): “The net value of an individual’s planned trades is identically zero”. According to them, if a person plans to buy two goods $x$ and $y$ in period 1 with money $S_{m,0}$ received in period 0, then:

$$P_x d_x + P_y d_y - S_{m,0} \equiv 0$$

In an extension of Say’s Principle, the value of an individual’s planned excess demand (demand less supply) for the $m - 1$ variety of goods plus his planned excess demand for money will be identically zero:

$$P_1(d_1 - S_{1,0}) + P_2(d_2 - S_{2,0}) + \cdots + P_{m-1}(d_{m-1} - S_{m-1,0}) + (d_m - S_{m,0}) \equiv 0,$$

or

$$P_1 x_1 + P_2 x_2 + \cdots + P_{m-1} x_{m-1} + x_m \equiv 0 \text{ or } \sum_{i=1}^{m} P_i x_i \equiv 0 \text{ (} x \text{ stands for planned demand)}$$

In an economy of $k$ individuals, the value of aggregate excess demand for good $i$ is:

$$\sum_{j=1}^{k} p_i x_{ij} = P_i \sum_{j=1}^{k} x_{ij} = P_i x_i,$$

so the aggregative version of Say’s Principle is also $\sum_{i=1}^{m} P_i x_i \equiv 0$.

Clower and Leionhufrud (1973[1981], p.92) thought Say’s Principle could explain excess supply (ES) of commodities, especially persistent large-scale unemployment:

> It is impossible for all trade to be executed as planned, so prices and trading plans must be revised. Some of the commodities in aggregate ES may be labor services, so SP is consistent with existence of a large scale unemployment.

> SP is also consistent with indefinite persistence of unemployment on a large scale, for it involves no assumptions and yields no implication about the dynamic adjustment behavior of the economic system.

Clower and Leionhufrud (1973[1981], p.95) also believed that SP could explain oversupply of all currently produced commodities:

> Disequilibria in which the sum of the values of EDs for all currently produced commodities is negative and equal in value to the positive ED for money. This means on balance, that the entire business sector is under general deflationary pressure. The typical industry will be laying off workers. If there are some industries hiring, they won’t hire enough; unemployment will be widespread.

The proposition of Clower and Leionhufrud is very similar to Walras’ Law. The main difference is that Clower et al introduce “planned trades” instead of the real supply and demand in market. They also criticise that the latter cannot explain individuals’ decision-making behaviour. However, there are some problems with their inference.

In the first place, planned price (estimated price) may not be same for each individual. If so, it is impossible to formulate out price $P_i$ in the ag-
aggregate form of \( SP \), namely: 
\[
\sum_{j=1}^{k} P_{ij} x_{ij} \neq P \sum_{j=1}^{k} x_{ij}.
\]
So the aggregate version of \( SP \) should be: 
\[
\sum_{i=1}^{n} \sum_{j=1}^{k} P_{ij} x_{ij} = 0.
\]
Thus, the left hand side of the equation is no longer the value of aggregate demand.

Second, as they realised, it is unlikely for all trades to be executed as planned. Since the planned excess of demand is not equal to the real one in the market, the real situation in the market cannot be explained by ‘planned trades’. For example, since individuals can adjust their supply and demand according to a changed price in the markets, it is possible that all goods achieve market equilibrium even where there is excess demand in the planned trades.

Third, as was stated previously, an individual cannot plan all the goods that might be demanded. Thus, attempting to theorise based on assumptions concerning planned trade cannot apply to every individual, and to every good and service.

4. BUSINESS CYCLE THEORIES IMPLIED BY SAY’S LAW

The late eighteenth and early nineteenth centuries witnessed the first Industrial Revolution in Europe. The use of machinery considerably increased production capacity. “Physiocrats” worried about the possibility of overproduction and put forward the doctrine of “Consumption is the measure of reproduction”. In this climate, Say’s opposing law was born. So from the very beginning, Say’s Law was addressing the possibility of a business cycle problem. During the evolution of Say’s Law, more ideas were added about the business cycle. This section places these implicit business cycle theories into three categories: partial gluts, reverberation of partial gluts, and temporary gluts.

A. Partial gluts

It is a popular idea among Classic economists that economic recessions are due to partial gluts and general gluts are totally impossible. There are many explanations for this claim, among which James Mill’s is typical. In his book, Mill (1808, p84-85) wrote:

It may be necessary, however, to remark, that a nation may easily have more than enough of any one commodity, though she can never have more than enough commodities in general. The quantity of any one commodity may easily be carried beyond its due proportion; but by that very circumstance it is implied that some other commodity is not provided in sufficient proportion. What indeed is meant by a commodity’s exceeding the market? Is it not that there is a portion of it for which there is nothing that can be had in exchange? But of those other things then the proportion is too small. A part of the means of production which has been applied to the preparation of this superabundant commodity should have been applied to the preparation of those other commodities until the balance between them had been established. Whenever this balance is
properly preserved, there can be no superfluity of commodities, none for which a market will not be ready.

According to this theory, economic recessions are caused by partial gluts due to coordinate failure, disproportion, or miscalculation. By partial gluts, this theory suggests that, if there is an excess supply of one good, then there must be an excess demand for other goods, and the total value of excess supply will equal the total value of excess demand. This theory both contradicts itself, and doesn’t match with the reality.

First of all, partial gluts are theoretically impossible if it is claimed that excess supply is accompanied by excess demand. The foundation of this theory is that demand means purchasing power. By this definition, if some goods fail to be sold out (oversupply), the expected income and thus purchasing power will not be fully achieved, and thus the effective demand will be less than the planned demand. As a result, there is no (effective) excess demand in the market at all. Consider an example economy of two individuals (1, 2) and two products (food and clothes). For convenience, assume the price of both the food (F) and the clothes (C) are $1, and it begins in equilibrium: individual 1 produces 8F but demands 3F and 5C, while individual 2 produces 10C but wants 5F and 5C. Suppose individual 1 mistakenly thinks that the demand for clothes is high, so he produces 6F and 2C. In the market there will be an excess supply of 2C because individual 1 only wants to buy 3C but individual 2 has produced 5C, but there is no excess demand for food because individual 2 now has only 3 dollars to buy 3F due to his unsuccessful selling. Thus if an individual’s supply is not fulfilled, his planned demand is not effective, and there cannot be any excess demand to match any excess supply. If so, a glut of any commodity implies an overall glut in all markets, namely, general gluts.

Second, in reality partial gluts are possible, but are unlikely and short-lived. With the existence of savings and inventory, both excess demand and partial gluts are possible. On the one hand, saved money can be used to buy whatever is wanted so an individual’s excess demand is effective even if his supply is not successful (partial glut). On the other hand, inventories in industry will decrease where there is excess demand in a market, so the excess demand is perceptible.

Nevertheless, this situation is unlikely to occur in reality. One reason for this is that price is the indicator of market supply and demand, and will automatically adjust to changes. If demand is greater than supply, the price will increase, which will attract more suppliers. If supply is greater than demand, the price will drop and producers will supply less. Another reason is that entrepreneurs are very sensitive to inventory and price changes. They adjust their supply according to the market demand, so a massive oversupply of one variety of good is unlikely. Even if partial gluts do happen due to some mistakes or miscalculations by entrepreneurs, they will be very
short-lived: the high profits in a high demand industry will attract funds and producers in oversupplied industries with excess supply, so any excess demand or supply will be quickly eliminated.

Third, if demand simply means desire to consume, excess demand is possible, but it will not rule out the possibility of general gluts. Mill (1808, p.85) was right to say that oversupply of a commodity means "there is a portion of it for which there is nothing that can be had in exchange", but his attributing the oversupply to disproportion is not rigorous. A shortage of a commodity may occur due to misapplication of the means of production, or due to an inability to produce it because of technology constraints. In the former case, according to Mill, it is a partial glut situation. However, for the latter case, as discussed previously, since people keep some as yet impossible aspirations in mind and do keep money aside for the prospect of new products or services, general gluts may occur.

Fourth, this theory underestimates the function of the market mechanism. Although market failures may occur due to information asymmetry, free goods, or externalities, facts show that the market mechanism is a very efficient way to reallocate resources. The planned economy in some communist countries such as the former USSR and China was proven to be an expensive mistake. There are many suggestions for the cause of partial gluts, such as misdirection, misapplication, miscalculation, disproportion, accumulated errors in production and coordination failure. All these can be put into the category of market failure. As discussed above, with the adjustment of price in a competitive economy, market failures cannot persist in multiple sectors for extended periods of time. Even in a monopoly economy, monopolies can detect the potential overproduction of commodities from increases in inventory and can adjust production accordingly. In short, the invisible hand of the market mechanism works better than the most competent central planning.

B. Reverberation of partial gluts

Partial gluts cannot explain the fact that overproduction occurs in so many industries during periods of economic recession. To explain widespread gluts in an economy, Jonsson (1997) produced his reverberation theory.

Based on Clower and Leionhufrud’s interpretation of Say’s Law, Jonsson introduces another concept: attempt trades. At the outset, Jonsson (1997, p.208) emphasises that “in any market which does not clear, the short side always dominates and rations quantities to a smaller number than that specified by equilibrium”. If an individual fails to sell what he had planned, the planned income is not realised. So he may be forced to revise his ex ante consumption plans. Jonsson demonstrates the effects of partial gluts in two ways. First, unemployed workers may attempt to buy fewer consumer goods than they would if they were employed, which would reduce
the demand for consumer goods. In this way, gluts in some areas will exacerbate gluts in others. Second, when suppliers fail to sell all that they had planned, they may (1) try to sell more of something else in order to compensate for the shortfall of income and (2) constrain their demand for the things that they had intended to buy. That is they will attempt to sell more and demand less. As a result, the attempt trades are not balanced: \[ \sum_{i=1}^{n} \sum_{j=1}^{k} P_{ij} d_{ij} < \sum_{i=1}^{n} \sum_{j=1}^{k} P_{ij} s_{ij} \] or \[ \sum_{i=1}^{n} \sum_{j=1}^{k} P_{ij} x_{ij} < 0 \] (here \( d \), \( s \) and \( x \) stand for attempt demand, attempt supply and attempt excess demand respectively). Jonsson concluded that the actual excess demand is negative for: \[ \sum_{i=1}^{n} P_i \ast ED_i = \sum_{i=1}^{n} \sum_{j=1}^{k} P_{ij} x_{ij} < 0. \]

In his theory, Jonsson notices and accepts two phenomena in economic recessions: widespread gluts over many sectors and massive unemployment. From these two facts, it can be concluded that economic recession is characterised by general gluts. If overproduction only occurs in a few industries (the production in these industries will then stagnate or decrease), the growth in other industries will compensate for this decline, and the economy will appear to still be growing. Thus general gluts, as opposed to partial gluts, are required for a recession. The existence of unemployment (together with unemployed capital) means idle production capability, which can only be explained by the general overproduction in the economy (if there is any industry with excess demand it will employ these idle factors). However, his reasoning about reverberation is flawed. We consider here three aspects.

First, it is true that unemployed workers will buy less consumer goods. But the existence of unemployed workers also means fewer products will be produced in the corresponding industries. The decreased consumption due to unemployment will be offset by the decreased output, leaving the value of total oversupply in the economy unchanged. Hence unemployment will not exacerbate gluts.

Second, it is reasonable to assume that in an economic recession, people will try to sell more and buy less, but this will not affect the attempt trades balance. On one hand, people want to sell more because they want to increase their income so that they can consume more. So, if they successfully sell more they will also consume more of other goods in the same value. This will not alter the trade balance. On the other hand, people demand less because of their reduced income from unsuccessful selling, so the value of reduced demand will always equal that of their unsold goods and the total excess supply will not be changed.

Third, attempt trade in this theory acts as a revised version of the previously disproved planned trade concept. Although it is closer to real trade than ex-ante planned trade, it cannot explain the real situation in markets. If an individual’s attempt supply is successful, the attempt demand will match the attempt supply and thus there is no excess demand in the
market. Otherwise, if an individual’s attempt supply is unsuccessful, the attempt demand will not be effective. So, excess demand in the market is a different form to that of attempt trade.

Although Jonsson’s explanation is questionable, his suggestion that partial gluts may result in general gluts \( \sum_{i=1}^{n} P_i \times ED_i = \sum_{i=1}^{n} \sum_{j=1}^{k} P_{ij} \times x_{ij} < 0 \) is a useful concept. Actually, partial gluts may lead to general gluts via a different route. Overproduction in one industry may result in capitalists seeking other profitable ventures. As a result, capital and labour will shift to other industries and after a period of time gluts may occur in these industries. These processes will continue as long as there are any profitable industries. Eventually, when all the industries (or most industries, or the major industries) become unprofitable, an economic recession will occur.

**C. Temporary gluts**

This theory is implied by Say’s equality. As Baumol (1977, p.159) claims, Say’s equality ensures that “supply and demand are always equated by a rapid and powerful equilibration mechanism”. Assuming Baumol’s claim is valid it may say that oversupply is only sustainable for short periods, or that Say’s equality allows temporary general gluts to occur.

Baumol (1977) does not describe what the “rapid and powerful equilibration mechanism” is which equilises demand and supply. But from his paper in 1952 can be seen that a change in cash balance can affect commodity prices. According to Baumol (1952), in the classical system modified by Patinkin (1948, 1949, 1951), a decrease in money supply or an increase in money demand from equilibrium will lead to excess demand in the money market, and hence excess supply in the goods market. The excess supply of goods will cause the price level to fall. As the price level falls, the money demand will decrease. The price level will fall to a point where the excess demand for money or the excess supply of goods is eliminated. Thus, the money market and the goods market will automatically restore equilibrium.

The reasoning appears to be correct, but the problem is whether the fall of the price level can always eliminate the excess money demand. To start with, consider an extreme case: a liquidity trap. When people have lost confidence in the economy, the demand for money skyrockets and the price level of goods may fall to zero. Even so, the excess demand for money may not be eliminated. Second, the price level of goods cannot fall below the production cost (otherwise, it is pointless to produce them). So, it is possible that the fall in price will be insufficient to equalise demand and supply in money market and thus the money goods markets will not automatically restore equilibrium. Finally, the price level of goods is nothing but the relative price of the total goods in terms of the total money balance. Just as the change of relative price of goods may fail to remove partial gluts due to cost constraint, the change of price level may fail to eliminate general gluts of goods. In short, since the fall in price level may fail to restore
It is instructive to consider a few examples. According to National Bureau of Economic Research of USA (NBER), in the period between 1854 and 2001, the US economy experienced 32 business cycles, with an average length of each economic contraction being 17 months. During the period the longest economic recession lasted for 65 months (from October 1873 to March 1879). Even with the assistance of sophisticated monetary and fiscal policy, the average length of economic recessions in modern times (measured from 1945 to 2001) is 10 months. Such a long period of economic recession shows that the so-called “rapid and powerful equilibration mechanism” does not work well, and economic recessions cannot be simply explained by market disequilibrium.

Another flaw in the theories is the existence of unemployment and unemployed capital, which has become a widespread and persistent phenomenon. Baumol (1997) justifies the persistence of unemployment using the partial disequilibrium and the special attributes of the labor market. He is correct in stating that “we know that Say’s Law, even in its strongest variant, does not require each and every market to be in equilibrium” and that “labour is a commodity which cannot be increased and diminished at pleasure.” However, he overlooked an important fact that labour is a factor input for production. An oversupply in the factor market (e.g. unemployment) is due to overproduction in the goods market, so persistent unemployment implies a long period of general oversupply in the goods market, as does the idle capital. This oversupply leads to wasted factors which prevent the economy growing at its full speed. In this sense it could be said that today’s economy is caught in a low growth trap.

5. THE ESSENCE OF BUSINESS CYCLE

Some classic economists are wrong to attribute economic recession to partial gluts and to deny general gluts, but their attacks on Keynesian economists seem reasonable in that demand deficiency is hard to imagine and wasteful consumption is not a solution to economic growth. There seems to be a dilemma here: if general gluts are admitted economy-wide overproduction or demand deficiency must be accepted; if demand deficiency is denied it follows that partial gluts theory must be acknowledged. The solution to this dilemma is that economic recession can be associated with an oversupply of old products and a shortage of new varieties of goods. Due to innovation scarcity, factories fail to produce new types of products that consumers want to buy; instead they produce too many old products. As a result, gluts or underconsumption of old products spreads in the economy.
From this reasoning it can be concluded that the essence of the business cycle is innovation scarcity. Why then is innovation scarce? How does it affect the economy? These are questions we address in this section.

**A. Reasons for innovation scarcity**

Innovation means the introduction of novelties: creating new things or making changes to something already established. It includes technique innovation, management innovation, new products, new ideas, new designs, etc. Why is innovation scarce now? What deters innovation activities? Two reasons are obvious.

(a) High risk

Innovations are creative activities. To innovate is to do something that has never been done before, so innovators are not sure they will succeed and what they can do is to try again and again. Facing the possibility of innovation failure, risk-averse investors are reluctant to put their money on innovations; instead, they prefer to invest in production that has a relatively certain return. As a result, innovation investment, or R & D funds, is severely scarce, giving rise to innovation scarcity.

(b) Imitation

Another reason for the shortage of innovation is the imitation. Innovations require hard and intelligent work which takes a long time and a great deal of money, but imitation is very easy. For example, to copy the software which has cost 10 years and millions of dollars only takes a few minutes. So the externality of innovation is enormous. Just like public goods are under-invested due to the free-riding problem, innovation is underinvested and becomes scarce.

Fortunately, human beings invented Patent Law, which protects the benefit of innovators and encourages innovation in a certain sense. However, not all countries have patent laws, and many countries that do have, do not implement them well. Moreover, the present Patent Law is not good enough to solve the problem of scarce investment in innovation.

**B. Economic functions of innovation**

It is well known that innovation has positive effect on economy. This can be explained with reference to two economic functions of innovation. First, innovation can increase efficiency and reduce production costs. Release of cost constraint in production leads to lower prices in a competitive economy, and thus increases the levels of both supply and demand. As a result the economy expands. For example, technique innovation in the Japanese auto industry reduced costs and promoted economic revitalization after WWII.

Second, innovation creates new products and services, which provide new outlets for investment and expand aggregate demand. Because of consumption constraint, if no new products or services emerge, the total output will be limited to the constrained quantity in the long run, so living standard
will stagnate. Due to innovation, new products release the consumption constraint, and the economy keeps growing.

From these two functions of innovation it is apparent that innovation is the key to break cost and consumption constraints, and thus it is an essential factor determining the progress of economic development.

C. Innovation and business cycle

From foregone discussion, it is safe to say that innovation scarcity is the essential factor accounting for business cycle. In reality, there are some interactions between innovation and the business cycle.

(a) Innovation scarcity leads to business cycle

Due to the high risk of innovation, firms are reluctant to invest in innovation (unless forced to by competitors) and prefer to expand production that is relatively safe. The lack of innovation results in both cost and consumption constraints, which give rise to limited aggregate demand. When aggregate supply exceeds aggregate demand, the economy will begin to contract, and thus economic recession starts.

(b) Business cycle promotes innovation investment

When economic recession happens, firms find their production is not profitable, so they will do their best to find other projects or products that are profitable. When all profitable projects are exploited, firms will find themselves at a passé: No project is profitable even if there is no barrier for exit or entry. In this passé, firms have no choice but to invest more in innovation.

(c) Innovation ends business cycle

The increasing investment in R&D promotes innovation. As a result, hi-tech machines and new products will release the cost constraint or consumption constraint, and aggregate demand expands. Consequently, the economy recovers and expands, and the business cycle has been completed.

6. THE WAY TO REMOVE BUSINESS CYCLE: REVISING PATENT LAW

To encourage innovation, governments in many countries provide large amounts of R&D funds. However, governmental funds are limited, and in most countries governmental R&D funds are concentrated on defence. To solve the innovation scarcity problem, we should encourage the private sectors to invest in innovation. The most efficient and practical way is to revise Patent Law.

No one can deny the success of Patent Law, but it is also true that the present Patent Law has some shortcomings and thus needs revising. By revising Patent Law thoroughly, innovation investment may become popular and our economy and society may enter a new era.

A. Shortcomings of today’s Patent Law
There are some arguments about the negative effect of Patent Law and it still remains an open topic. However, since discussion on it is not the emphasis of this paper, we only provide the following brief opinion.

(a) The methodology of patent system

The patent system can be traced back to medieval guild practices in Europe. To raise revenue, monarchs in European countries frequently sold their privileges such as monopoly over trade in specified commodities. According to Meinhardt (1946, pp.30), in Britain, “ever since the 12th or 13th century the Crown, in the exercise of the Royal Prerogative, granted charters and patents conferring monopolies on trade guilds, corporations and in some cases on individual persons.” The first general statute about monopoly grants to innovators is often thought to have been passed by the Venetian State in 1474 (Reid, 1993). The principle that monopoly should be granted only for innovators was laid down by Francis Bacon in 1602, and it was adopted by the British Parliament in 1623. The Statute of Monopolies, Section Six, authorised:

Letters patent and grants of privilege for the term of fourteen years or under, hereinafter to be made, of the sole working or making of any manner of new manufactures within this Realm, to the true and first inventor of such manufactures, which others at the time of making such letters patent shall not use.

Although there are different expressions in the patent laws of different countries and many reforms of patent law have been made worldwide, Section Six of the Statute of Monopolies is still the foundation of present patent laws.

The patent system is very successful, but has also caused some problems, among them the abuse of patent monopoly, such as the wrongful suppression of patented inventions, charging excessive prices for patented articles, and unwarranted threats under patents. Many efforts have been made to counteract the abuse of patent monopoly. Although some are effective, because of the complex procedure and high cost of these efforts, they complicate patent law and lead to an inefficient society.

In the present post-industrial age the patent monopoly seems neither desirable nor necessary. On one hand, patent monopoly, like any other monopoly, is not in harmony with democracy and market economy. One who has a monopoly power may want to profit at the expense of the community. On the other hand, unlike in traditional times there are numerous companies and abundant funds in society today, so inventors can easily profit by selling patent technology or granting licences under a patent. Thus, it is not necessary to give the power of a monopoly to an inventor.

Such a power, being unwanted and unnecessary there is no reason to maintain it. The replacement of the patent system with a licence or copy-
right system is a better choice for society and there would be few disadvantages for inventors.

(b) The duration of patent right

The duration of patent right varies from country to country and depends on different types of inventions, but the maximum to date is 20 years. This time limitation frustrates inventors in more than one way. First, the life span of some innovations is much longer than 20 years, so the duration of patent right partly reduces the return to inventors. Second, some innovations may require very high levels of funding so innovators may not make a positive return in the limited time available. Third, industrialisation of some patent technology may take a long time, so the protection provided by patent law is very limited. For example, in the medication industry, before commercial production of a new medicine, testing takes many years by law. As a result, when firms start to sell the new medicine, they are no longer protected by patent law.

Based on the above reasons, it can be confidently asserted that the duration of patent right should be prolonged in modern society. To be fair to innovators, it should be infinite. Innovations are the products of innovators, so they are the property of innovators. Just as any other physical property belongs to the owner until it is sold, patent rights should belong to innovators without a time limit.

It may be argued that the duration of the patent right is necessary to fully implement the patent technology: innovators are normally anxious to apply their patents before the time limit expires. The reasoning sounds logical, but it is on the wrong track. In reality, many companies wait for the expiration of patent technology in order to avoid paying the patent fee, so the time limitation actually delays the implementation of patent technology. Moreover, the application of patent technology is not the responsibility of innovators, so innovators should not be punished by the imposing of a time limit. The concept of the duration of patent right seems as foolish as fining farmers if they have not sold their crops within a certain period of time.

(c) The protection scope of Patent Law

Present Patent Law only protects practical inventions. Neither Patent Law nor Copyright Law protects scientific discoveries and theoretical breakthroughs. As a result, little private funds support theoretical and scientific research. In fact, theoretical work is very important and largely determines the pace of practical inventions, for they are the base and the source of invention. To guarantee the innovation speed, something must be done to encourage scientific and theoretical research.

(d) The compulsory licence rule

The compulsory licence rule was adopted to counteract the abuse of patent monopoly. It stipulates that, on the expiry of 4 years from the date
of the application or on the expiry of 3 years from the date of the grant of the patent right, if the patent has not been exploited or has not been exploited fully without any justified reason, the patent office in the member country may grant a compulsory licence to exploit the patent upon request. Once again, this rule punishes innovators for not applying the innovation in time and discourages innovation activities.

B. The impact of revising Patent Law

If the patent law was revised thoroughly — abolishing monopoly, prolonging the duration infinitely, fully covering all innovations and abandoning the compulsory licence rule, its impact on the economy and society would be significant.

(a) Improved economic growth

Since the benefits of innovators would be fully protected, innovation scarcity may be mitigated or even removed. At least, the balance between investment in production and in innovation would be improved, so overproduction could be avoided. As a result, major economic fluctuations would be obsolete, the economy will develop faster, and living standards of mankind would improve.

(b) Improved economic environment

The removal of innovation scarcity would improve the economic environment in two ways. First, no large economic fluctuation means that the invisible hands will work well and government intervention will be minimized. Second, innovation would create substitutions for old products, so monopolies in old industries would collapse. Thus, with revised patent law, the economy is likely to become more competitive and more efficient.

(c) Changed social structure

In an industrial society it is the capitalists who are very powerful because capital is the scarcest resource. Under a new patent law, innovation would become the most precious resource, and innovators would be the richest and most important class. Everyone respects knowledge, so a knowledge society would replace the industrial society, and the so-called “new economy” or “knowledge economy” would begin to emerge.

7. CONCLUSION

The validity of Say’s Law depends on whether the will to purchase falls short of the power to purchase, and consequentially depends on if there are sufficient new variety of products in the goods market. Due to high risk of innovation activities and the imitation problem with new products, innovation scarcity is a significant problem. As a result, business cycles occur recurrently. If innovation scarcity is removed by revising the present patent law thoroughly, the business cycle may disappear and human being would enjoy rapid and sustained improvement in living standards.
REFERENCES


http://www.nber.org/cycles.html/


