Classical Inflation Model with Demand Pull and Supply Resistance

Artigo apresentado ao XIX Encontro Nacional de Economia Política (Março – 2014)

Alberto Handfas

Universidade Federal de São Paulo

New School for Social Research

Abstract

This article discusses an alternate theory of inflation by resuming the classical political economy debate on accumulation and money. A Marxian money theory, updated to a fiat-money, credit-driven economy is then combined with a Smith/Ricardo/Marx dynamics where the rate of profit is the limit to growth, rather than the neoclassical full employment. These two distinguished approaches together give rise to a demand-pull, supply-resistance inflation model in which the interaction of injections of purchasing power with the proportion of profit recommitted to production is the main explanatory variable. The unbalanced multi-sector dynamics of an economy allows such relationship to yield a usually nonlinear behavior.

Resumo

Este artigo discute uma teoria alternativa de inflação retomando o debate da economia política clássica sobre acumulação e moeda. Uma teoria marxiana da moeda, atualizada a economia de moeda escritural e de crédito, é então combinada com uma dinâmica de Smith/Ricardo/Marx onde a taxa de lucro é o limite ao crescimento, ao invés do pleno emprego neoclássico. Essas duas abordagens juntas levam a construção de um modelo "*demand-pull, supply-resistance*" de inflação no qual a interação de injeções de poder de compra com a proporção de lucro recomprometido à produção é a principal variável explicatória. A dinâmica multi-setorial desbalanceada de uma economia leva a não-linearidade na função de inflação.

1. Introduction

This article intends to discuss an alternate explanation for inflation based on the classical political economy debate on money and on capital accumulation. Departing from the difficulties conventional theories of inflation have had to reconcile with challenging historical events – such as stagflation, or low-growth/high-inflationary (and vice-versa) processes, it is presented here a Classical/Marxian inspired contribution (Shaikh, 1999 and 2012, Handfas, 2012). Pressures on general prices are to be explained by the interplay of demand-pull and supply-resistance forces, where former and latter are respectively endogenous-money and rate-of-profit driven. Elaborating on such framework and unfolding its arguments, the aim of this work is twofold: to present an empirically testable classical model of inflation and to carefully explain its functional form on theoretical grounds.

We aim to shed light on the way the interplay capital expansion hindrances and excess demand impacts inflation. With this in mind, and elaborating on the classical political economy debate about money and growth, we present here a **demand-pull/supply-resistance** model of inflation. Inflationary pressures are driven therefore by the combination of injections of purchasing power into the commodity market (a gauge for the demand pull) and the capitalization ratio – i.e. the percentage out of profits being reinvested in production --, (a proper gauge for the supply resistance).

We call the latter gauge the "***throughput coefficient***" (Shaikh, 1999). Elaborating on Marx's schemes of expanding reproduction, one can see that the maximum growth rate is reached when all surplus value is reinvested; At this point, such growth rate equals the rate of profit. The theoretical meaning, structural dynamics and consequences on inflation *of the* ***throughput coefficient***, "τ",will be elaborated further over the next sections.

As it follows, the classical perspective we intend to pursue here leads us to a breaking out of the straightjacket assumptions usually found in most traditional inflation theories. One can summarize those conventional assumptions in three main orthodox-mainstream hypotheses (i) the neutrality-exogeneity of money; (ii) the full utilization of productive factors and (iii) the steady-state, balanced, growth (which also includes here most traverse analysis theories).

Instead, the classical approach to the problem of growth limit in potential supply could well be traced back through Marx's explanation – which was later re-discovered by von Neumann: "the maximum growth rate in any self-reproducing system is equal to the profit rate" (Kurz, 1995, pp. 383-384; quoted from Shaikh, 2014, pp. 67).

There is no reason though to take for granted a linear and straightforward relationship between these two variables and inflation, as the third orthodox assumption would lead us. Because real growth is always unbalanced, capital accumulation brings about a turbulent and very frequently uneven system in which different industry branches might grow at different rates. In order to better investigate such sequence of events, we go over a Sraffian multi-sector framework of commodity production. Elaborating on them we find out a possible mechanism that would explain the throughput nonlinear impact on prices. We then mathematically derive from it a quadratic function representing such relationship to be empirically tested for inflationary as well as deflationary environments – and for credit expansion as well as retraction trends.

2. The Theoretical Conundrum

Two main approaches have dominated theories of inflation in traditional economics. On the one hand the quantity theory type of story emphasizes the long-run neutrality of money supply over output. More recent versions of this approach – as the monetarist/Nairu one – also underlie the explanation of such neutrality on the grounds of full employment. Most Keynesian types of theory, on the other hand, would counter the idea of neutrality precisely by arguing that the economy is not always in full employment. Nevertheless, they both share the same perception that, when in (or close to) full employment, output is not responsive toward an increase in money supply and that only prices are affected by it. A general representation of this stance would be:

$$\pi \approx f\left(\frac{\overset{+}{y}}{y_{\max}}\right) \qquad (1.0)$$

$$\text{or } \pi \approx f\left(\frac{\overset{-}{u}}{u_{\min}}\right) \qquad (1.1)$$

where π is inflation and y, y_{max}, u and u_{min} are respectively the current and maximum output and the current and minimum unemployment (the "natural" rate or the Nairu, depending on the version of the story)[1]. While Keynesians generally accept the persistence of a gap between y ad y_{max} or u and u_{min} , monetarists would suggest the opposite.

In this sense both traditions in one or another way blame supply (factors-scarcity driven) rigidities for causing inflation. Any increase in money supply due to expansion of credit, fiscal deficit, or increased production of gold (or printing money) would boost demand, which outgrows such rigid supply -- leading to inflation. We will see that under a classical perspective, an alternate theory can be unfolded. Accordingly we present in the next two sections a different standpoint both on the theory of money to explain the demand-pull aspect of inflation and on the supply-resistance behavior.

3. The Demand-Pull

Back to the quantity theory of money, from Hume's specie-flow, to Ricardo's, to modern monetarist versions, they generally state that the sum of prices of all commodities produced in a period of time equal the number of coins in circulation times the velocity of their circulation. This is to say,

$$PY = Mv$$

where P, Y, M and v stand respectively for price level, output, money supply and velocity of money circulation. Also their respective growth rates are as follows: and a=0 (for simplification's sake v is usually assumed exogenous). From this one gets

$$p \approx f\left(\frac{u}{u_{min}}\right) \quad (1.2)$$

3.1. Endogenous Money

But without establishing a causation direction, both relations are no more than tautological identities. In the XIX century political economy debate, Ricardo would take the quantity theory argument to say that an increase in M would cause an increase in prices only. Therefore inflation would be caused by every increase in money supply greater than increases in production of all other commodities. (Green, 1992. pp 15) Marx, following here Tooke, would instead suggest a different picture that flows naturally from both the labor theory of value and the particularity of money as the general equivalent

[1] (1.0) and (1.1) would basically mean the same, granted y and u being inversely determined by certain observed relationship such as the Okun Law.

(Shaikh, 1979b, p. 31). Marx asserts that the above equivalence shows a different causality as well as a much more complex money content and role in its interplay with prices and production.

Firstly, his argument goes, the causation direction is the other way around: it is values created in the production sphere, i.e. (labor) value multiplied by the quantity of produced commodities, that determine money prices, PY, and these subsequently determine the quantity of money (gold coins for example), M_G, required for commodity exchange in the circulation sphere. So increases in money supply are a consequence rather than a cause: **money is endogenous**. The same rationale can easily be extended to any token money economy (Marx, 1976; pp. 123). And, also by extension, in a fiat-money economy, banks (including the central bank) passively create money, responding to the needs of circulation – to the growth of aggregate demand.

Secondly, changes in supply of money (gold) are different from changes in supply of any other commodity for the former does not have to be sold to be exchanged by itself. This implies that while an excess supply of any ordinary commodity drives its market price down, excess supply of gold becomes redundant in circulation and takes at first the form of idle coin building up reserve hoards (Shaikh, 1979b). Such non-circulating money is more than allowed; it is indeed required by any commodity production society. This is because erratic behavior of demand/supply and of price oscillations makes commodity owners to be preemptive and even parsimonious. The system as whole needs to have at its disposal a reserve of money, which has been withdrawn from immediate circulation, to be used according to market fluctuations and its prospective requirements of coins for circulation. This has become even truer with the development of modern banking and financial markets and their greatly speculative behavior. From this, one should take into account that only part of the new supply of money will really go to the commodity market (to the sphere of circulation) and count as a potential pressure over prices (Marx, 1976, pp. 134). Roughly the same point was also made by Keynes, who would argue that "entrepreneurs will direct money to production if they expect that to be profitable (in terms of money)." Otherwise, "money is kept idle" (Keynes, 1937).

Thirdly, even that part of "active" money, which goes to circulation, will not necessarily put pressure over prices by increasing aggregate demand over supply. Recall that Marx distinguishes between money as money and money as capital (Marx, 1976; pp. 247-257). Money is not only used for buying commodities but also for investing and expanding their production. An increase in Money supply has therefore a dual effect. If it generates both more demand for and more supply of commodities roughly in the same proportion it will keep eventually prices unchanged. Any momentarily increase in effective demand caused by such excess of money will no more than temporarily increase prices of some commodities that will have their profitability increased and thus attract investments leading to an expansion of production. This in turn brings their prices down back to the original level[2].

Fourthly, Marx had a particular interpretation for movements of the general price level. Considering either a commodity-money or a token-money type of economy (as the world used to be in his time), the long-run price level is determined by both: (i) the (weighted average) price of the commodities (or of a representative bundle of them) relative to silver or gold (the commodity chosen to be money); and (ii) the monetary price of gold (or silver). While (i) depends upon structural conditions – driven by the endless process of equalization of profit rate between gold's and other commodities' industries (considering here each industry rebalances in supply and demand, costs, productivity and relative prices) --; (ii) is given by the exchange rate between gold and the country currency.

In a convertible token-money economy, that exchange rate is fixed and thus prices would ultimately depend on (i). Conversely, in an unconvertible money economy, that rate freely floats and it is its fluctuation that would eventually drive the general price level. In this latter case thus prices would more directly reflect the impact of macroeconomic conditions on the money-price of gold. Thus increases in macro variables such as excess demand, for instance, may lead to an increase in the quantity of token-moneys in a country, which everything else constant would erode the gold content of a unit of such token-money, raising then the money prices of commodities (whose golden-prices are still unchanged).

Marx did not talk openly on what general price movements would look like in an all-fiat-

---

[2] Keynes here would also agree that if part of the increasing money supply were to be used for new investments, it would increase supply of commodities without boosting prices – his notion of investment demand for money (Keynes, 1937b).

money economy although he referred *en-passant* to it when commenting on a few exceptional historical events. Yet, it is possible to derive it from his previous reasoning on unconvertible-money price determination (Shaikh, 2012; Ch. 5). In a fiat-money economy, those unconvertible tokens are no longer linked to gold -- and gold is not even an indirect medium of prices anymore. Thus the difference from the unconvertible story is that, **with fiat-money, gold no longer intermediate the macroeconomic conditions** (i.e. excess demand, economic activity, growth limit etc). The latter would impact directly commodity prices (Shaikh, 2012; pp. 60-68). All other Marx's premises (first to third) are still valid and in place regardless though: money is still endogenous and, when it goes to the commodity market, might have a dual character (boosting demand but also supply).

In the last almost a century, economies have transited more and more toward an all-fiat-money system (whose mechanisms were already being well developed way before Nixon's end of the Gold Standard in 1971). Inflation has become a major economic issue precisely during this period. Therefore we intend to focus our attention on what would be a classical inflation theory more particularly for a fiat-money economy.

It follows from the four premises above (particularly from the last one) that a classical/marxian long-run inflation can be traced down by looking at two major macro conditions: excess demand and supply resistance..

Inflation (or deflation) hence will only occur in the presence of persistent mismatches between aggregate demand and supply. As a matter of fact, capitalist mode of production (self) generates mismatches between expenditures and sales, unforeseen expenses, replacement of fixed capital, expansion of output, offsetting price fluctuations, payment of dividends etc.

3.2. Persistent Excess Demand and Injections of Purchasing Power

Capitalists' needs to invest drive changes in deficit spending, which usually is to be financed by borrowing from stocks or borrowing from others. At a first level of abstraction -- for argument's sake – when presenting the expanded reproduction Marx assumes no bank loans (Marx, 1978, p. 568). In a modern, fiat-money, economy new loans (by private banks) and high-powered money (by central bank) are also created in the form of injections of purchasing power to respond such capitalists' needs to invest.

When such injections do not meet (going far beyond) those needs to invest they may lead to persistent episodes of excess of aggregate demand. Only then inflation is triggered. The opposite would also be true for deflation.

From that reason we shall no longer interpret the term $(\hat{M} - \hat{Y})$ of (1.2) as the increase in money above the increase in commodities. Instead, it is to be understood as the increase in that part of money that goes to the circulation sphere but cannot be used (it is in excess therefore) for the expansion of commodity production. In other words it is the persistent episode of excess demand, or the new injections of purchasing power (or credit to commodity market) that are not used for expansion of supply due to lack of its further expansion.

3.3. Excess Demand and its Credit Proxy

Once we have outlined such close conceptual relationship between excess demand and credit to commodity market, it is interesting to formalize their interplay for the economy as a whole. First we need to aggregate all sources of demand and supply in the national accounts. More formally one can say that in the demand side,

$$D = C + N + I + G + EX \qquad (2.1)$$

Where, D, C, N, I, G, EX stand respectively for aggregate demand, consumption demand, intermediate input demand, investment demand, government demand, and export demand. On the supply side,

$$S = Y + N + IM \qquad (2.2)$$

Where S, Y and IM are respectively aggregate gross output, net output and import supply (we assume here that intermediate goods supply equals demand). Then it is possible to define excess demand as follows:

$$E = D - S = (C + I + G + EX - IM) - Y \qquad (2.3)$$

Rearranging this would give us

$$E = [(C + I + G) - Y] + [EX - IM]$$

Based on the rationale presented in the previous section we can proxy the first term of the right-hand side of the equation above through New Domestic Credit. That would allow us to empirically estimate the excess demand. So that we get:

$$e \equiv \frac{E}{Y} = \frac{[(C + I + G) - Y]}{Y} + \frac{(EX - IM)}{Y} \approx \frac{\Delta \text{New Credit}}{Y} + \frac{(EX - IM)}{Y} \qquad (2.4)$$

Equation (2.4) shows us a way to empirically tackle relative excess demand as the sum of relative new domestic credit and the relative trade balance. Summing up,

$$e \approx \frac{\Delta\text{New Credit}}{Y} + \frac{(EX - IM)}{Y} \qquad (2.5)$$

Among all possible injections of purchasing power, "e", there are three main domestic sources: banks, which create new money when they extend loans; direct consumer credit; and the portion of government deficits financed by direct printing money or Central Bank credit -- indirect printing money (Filho, 2001). International trade (or more generally all transactions in current account) surpluses are also a source of inflows (Shaikh, 1999; Kalecki, 1954) -- when exports outgrow imports, part of the aggregate supply has been sent out increasing the excess demand. In short, e = New Credit to Commodity Market + Net Foreign Demand.

3.4. Resilient Excess Demand and Growth Hindrances

Now, back to the previous section re-interpretation of the inflation relationship suggested in equation (1.2). To start with, recall that in traditional economics, the ex-post identity of (2.1) and (2.2) is a requirement for long run, steady state, equilibrium. Excess demand, E, in this case is transient, for it is the fuel for new investments, which in the next period will allow supply to catch up with demand. Also in this case, therefore, it does not cause inflation -- and thus E, by itself, should not be confused with inflationary excessive money supply of equation (1.2).

But in a real world economy, such catching up need not happen either smoothly or automatically. Hindrances to growth in one or in many industries may delay or even prevent such ex-post identity for many periods and utterly disrupt any possible smooth (traditional economics) growth path. It is only when those hindrances show up (and stick around persistently), that E becomes inflationary.

So the excessive money in the commodity market capable of pressuring prices, as expressed in equation (1.2), only appears when excess demand **meets** a supply that is reaching its growth limit; i.e. when the current economy growth rate, g, is reaching its maximum rate, g_{max} imposing some resistance for supply to catch up with demand in the next periods. From this, and using the relation (2.5), we can re-interpret equation (1.2) as:

$$\hat{P} = \pi = f\left[e, \frac{g}{g_{max}}\right] \qquad (3)$$

where π, and g/g_{max} stand for inflation rate, and a growth limit gauge. Equation (3) shows that inflation is a function of (is caused by) both, demand pull and supply resistance.

Our next step is identifying how those underlying variables interact in the process of creating such episodes of persistent excess demand/supply resistance, as well as grasping the mechanisms that transform them into inflation.

4. The Supply Resistance

As seen in equation (1), traditional economics in its either Keynesian or Neoclassical versions have full employment of factors (labor and/or capital) as its barometer. Their inflation theoretical framework, as in the many versions of the augmented Phillips curve -- with Nairu, or capacity utilization, or current-to-normal output gap benchmarks -- flows from it although in different ways. Hawkish Neoclassical theorists along with the quantity theory/monetarist tradition are straighter to the point in this respect: The economy is naturally and usually in full employment, meaning supply of goods cannot be increased in the short-to-mid run. Thus any increase in money supply beyond the needed for the purchasing of supplied goods will end up raising effective demand and might swell prices without producing any effect in the real side of the economy.

Keynesians would counter that by insulating prices from demand via monopoly/oligopoly mark-up prices mechanisms so that increases in money supply will raise both output and employment levels. When the latter approaches full employment though the supply curve becomes inelastic (vertical) and prices start to go up. Ever since Phillips published his seminal paper (Phillips 1959) the long run relationship between gauges of economic growth limit and prices has become the most popular type of inflation models. A supposed tradeoff between inflation and unemployment was thought to be empirically proved -- it was, after all, observed for a couple of years in the post-war period. Later the idea yielded augmented Phillips curves in its different versions with different limits: capacity, natural unemployment Nairu etc. The infamous world-wide experience of the 1970's and 1980's when inflation and unemployment were both skyrocketing basically disproved such empirical claims, glooming both schools explanations on the topic.

4.1. The Throughput Coefficient

There is an alternative to both perspectives for explaining limits of growth following a classical standpoint. It does not incorporate the mainstream economics standard limits for growth, i.e. full employment of either labor or capital (capacity utilization). The former is ruled out on the grounds of the permanent creation of reserve army of labor – the "general law of accumulation" (Marx, 1976, p. 795). The latter should likewise be rejected since capital is self-expanding by definition and so is capacity. The boundary to growth is to be found elsewhere: in the limit of capital reproduction, or the "maximum expanded reproduction" (Shaikh, 1999, p. 99).

If investment is the part of surplus that is recommitted to production, then let

$$I = \tau P$$

where τ is the throughput coefficient, or the percentage of surplus, P, that will be reinvested in production – a.k.a. capitalization ratio. Also if output growth, g, is a positive function of capital accumulation, g_K , then

$$g = F(g_K) = F\left(\frac{I}{K}\right) = F\left(\frac{\tau P}{K}\right) = F(\tau r) \qquad (4)$$

where r is the rate of profit. As we will see next τ is a gauge of the "stiffness" of economic growth.

4.2. The Rate of Profit as the Limit of Growth

When Marx introduces the expanded reproduction schemes, he suggests as a hypothetical instance, a maximum recommitment to production. This happens when all surplus is reinvested and nothing is consumed by capitalists; i.e. τ = 100%. In this limit case, a maximum economic growth, g_{max}, would occur when capital accumulation, g_k, approaches its upper boundary:

$$g_{max} = f\left(g_{k,max}\right) = f(g_k|\tau = 1) = f\left(\frac{\tau P}{K}\right) = f\left(\frac{1P}{K}\right) = f(r) \qquad (4.1)$$

This shows that maximum output growth will ultimately depends on capital accumulation's upper boundary, the current rate of profit, which is in the last instance determined by the secular tendency for it to fall as well as by its counter-tendencies (Marx, 1978). This secularly dynamic process -- the (Kondratief) long waves -- is also complexly interconnected to shorter run cycles of tightening and loosening in potential accumulation and with the process of equalization of aggregate demand and supply due to market mismatches –Juglar and Kitchin cycles, so to speak.

Equation (4.1) also shows that there is a way to measure the degree of "utilization" of the intrinsic potential accumulation of an economy at a given period of time by simply comparing the current capital growth ratio, $g_{K,t}$, with the current rate of profit, r_t , (which is the maximum rate at which capitalists would be possibly willing to accumulate) by dividing the former by the latter. Such strain gauge of capital accumulation boils down to the "***throughput coefficient***" itself (Shaikh, 1999). In other words,

$$\text{supply-resistance gauge} \simeq \frac{g}{g_{\max}} = F\left(\frac{g_k}{g_{k,\max}}\right) = F\left(\frac{\frac{\tau P}{K}}{\frac{1P}{K}}\right) = F(\tau) \qquad (5)$$

Now putting (3), (4), (4.1) and (5) together:

$$\pi = F[\, e, F(\tau)\,] \qquad (6)$$

Next step in modeling such classical theory would require finding a functional form for (6) that is both economically meaningful and easily tackled econometrically saying. We already have a hunch that there should exist an interactive relation between "e" and $F(\tau)$, implying a multiplicative form such as

$$\pi = \theta_0 + \theta_1 e * F(\tau) \qquad (7)$$

where θ_0 , θ_1 are constant terms. That means changes in general prices will only show up when both $F(\tau)$ and excess demand differ from zero[3]. In any case, we still should investigate further the functional form of $F(\tau)$ itself – that is: the way through which the general (economy-wide) accumulation limit affects inflation. This is not trivial. Indeed, it would be unwise to mimic the way traditional inflation models relate their economic limit measures with change in prices since their limits considerably differ in content from the one we are suggesting here[4].

4.3. The Interplay of Credit Expansion and Tightening Accumulation

The impact of the tightening of capital expansion on the general prices is determined by a quite dynamic process – as opposed to the capacity/full employment static equilibrium idea. Nonetheless it might be possible to measure it by looking at the interplay of credit flows and our strain gauge function, $F(\tau)$, as expressed in equation (7). It allows us to grasp what happens when the new injections of credit (or purchasing power) enters the circulation sphere precisely when such straining accumulation is gaining momentum. It is

[3] External shocks, inflationary memory and other minor events are disregarded at this level of abstraction – they are all to be accounted by the constant α in equation (7).

[4] A comparison between different limit gauges is presented and discussed in further detail in Handfas, 2012.

here that the abovementioned persistent episodes of excess demand are set in motion.

For instance, assume that precisely during a downturn of the rate of profit cycle, a new bank credit (and/or trade surplus) inflow is leaking into the sphere of commodity circulation. Insofar as the lowering rate of profit makes capitalists lose their willingness/ability to invest, i.e. any further recommitment of the extracted surplus into production leads τ to soar, those new (money) injections are not met by the needed expansion of commodities output so that they are turned more and more into "means of circulation for circulation' sake". It is precisely and only at this point that such new credit to the commodity market creates inflationary pressures. If, on the contrary, τ is low, that new credit will instead be met with new expansion of production, which would neutralize those pressures by preventing persistent excess demand.

To use a Keynes cross type of analogy (a phase diagram representing current and next period output tradeoffs), a shift upwards in the aggregate demand curve would create excess demand (the vertical distance between the current output on that curve and the 45 degrees line) – setting in motion a path toward a new equilibrium with expansion of output greater than the original excess demand due to the multiplier. If output responds perfectly to excess demand no inflation is created eventually. But what if there are frictional forces in the middle of the path that would delay, constrain or even prevent such excess demand (planned production) to be straight forwardly transformed into new output? Those "drags" would represent the stiffness of growth. If they are in place, the friction between the injection of credit (excess demand) and its opponent drag would not only bother the smooth transformation of credit into new output – slowing down growth -- but also result in inflation[5]. Such drag can be measured by the ***throughput ratio***, the degree to which the growth of capital is approaching its limit.

We also shall recall that, for "structural" reasons, real world reproduction of capital is not an even process at all. Different sectors of the industrial chain experience different rates of accumulations and thus yield different inflationary, frictional, forces. They must impact inflation in a non-linear way. This gives us important hints on the functional form of the inflation model as we shall scrutinize next.

[5] In the same way that – mutatis mutandis – the friction between a body sliding over a rough surface would not only slow down the body trajectory but also generate heat (which is akin to inflation in our analogy).

5. Throughput and Bottlenecks in a Multi-sector Economy

Bottleneck formation is a well-known problem in mainstream growth theory. It shows up, in orthodox analysis, in the theoretical case of an economy that adjusts from one balanced growth rate to another -- it is therefore particularly tackled by non-equilibrium theories that have dwelt on the **traverse growth** issue (Hagelman, 1984, Amendola, 1984, Gehrke, 1997; Shaikh, 2014). However, traverse literature is generally about equilibrium paths, in which bottlenecks appear as exceptional and transient circumstances. It usually abstracts central aspects of real capital accumulation with its underlying transfers between sectors and price-profit anarchic responses (Shaikh, 2014). For this reason, instead of using this type of analysis in our inflation model we intend to foster a more real-world capitalist accumulation standpoint, in which growth is always unbalanced.

As a matter of fact, Marx was perhaps the first to unfold a type of "traverse analysis", when discussing the transition from simple to expanded reproduction -- in volume II of *Capital*. However, his schemes of reproduction were a theoretical/pedagogical device and never meant as descriptions of actual growth -- in which new technologies enter via new investment, as do new products. Accordingly he knew well that in real world growth, imbalances does not yield smooth transitions to another balanced path – as one can usually find in mainstream theories. Moreover, he -- as much as Keynes -- was particularly aware of the disruptions provoked by such imbalances on prices.

Erlich, while commenting on that, realizes that "Marx, in a striking passage that sounded astonishingly like Keynes's much quoted statement about 'pressure on the facilities for producing that type of capital [which] will cause its supply price to increase'.

> *"There is a check in reproduction and therefore in the flow of circulation. (...)The same phenomenon (and this as a rule precedes crises) can occur if the production of surplus capital takes place at a very rapid rate, and its retransformation into productive capital so increases the demand for all the elements of the latter that real production cannot keep pace, and consequently there is a rise in the prices of all commodities which enter into the formation of capital"*
>
> (Marx, Capital, Vol III p.313, cited in Erlich, 1967, p. 609)

Erlich unfolds a numerical example similar to a Marx's expanded reproduction scheme using a Harrod-Domar framework as the basis so that growth is driven by the savings rate. The transition to a higher growth rate requires a higher savings rate to finance the higher investment required. Higher investment first requires higher output in the sector that produces capital goods (department I). And that in the best conditions takes time -- 2 years in his example -- (Erlich, 1967, p. 614). Meanwhile the economy should respond through changes in prices.

5.1. Capital Goods Ceilings

So bottlenecks are cast by basic goods sectors. But more precisely, which of them tends to trigger the process? Erlich explains that "whenever an economy is experiencing a strong cyclical upswing, or whenever deliberate attempts are made to sharply lift the rate of growth from the hitherto prevailing level, the 'ceiling' in sector 1 area are likely to make themselves felt sooner or later, more likely than not – in industries with particularly high capital output ratios and long gestation periods." (Erlich, 1967, pp. 609, 610)[6].

When the growth rate is shifted up, there will be a sector under more "pressure on the [capital-producing] facilities". Investments and capital replacements will need to be "channeled toward" that low ceiling industry in order to enable it to increase production to be used as capital (inputs) for the other industries. Only then the economy is set onto a proper position for the new growth rate. The initial pressure is gradually relieved; but it is definitely only in a mid-to-long run framework (in Erlich's example, two year-long) that the adjustment can be fulfilled[7]. In the meantime, in the first stages of the adjustment, when the pressure is still not dissipated, it will express itself as inflation.

6.2. (Generalized) Basic Sectors' Ceiling

The same reasoning might be applied for any other demand-supply bottleneck, even if we generalize to a mutually interdependent system such as in one where all basic goods are heterogeneous, thus cannot self-sufficiently self-reproduce themselves and have different technical coefficients. In this case each and every (basic goods) industry depends on and produces inputs to the others (e.g. consumption goods are required for Sector I workers to consume so that consumer and capital goods are basic goods…).

Also we need not limit our investigation to the equipment producer bottlenecks. When delving into the hindrances found by capitalist economies to finance their initial stages of development, Kalecki went through the same problem and pointed out to bottlenecks in

---

[6] This should not depend on whether savings are endogenous (as in Marx) or exogenous (as in Harrod)

[7] Allowing for the introduction of new techniques (instead of the Harrod-Domar constant capital output) would not make a world difference in terms of the timeframe problem – for technical change is itself a time consuming process.

department II. When the economy is shifting its growth rate and boosting the industrial sector, new investments means increasing department I (capital goods) output. It also means department I will hire more workers who will need more department II (consumption) goods. If department II is operating in or around full capacity, he argues, consumption goods' demand outgrows supply. This leads to two intertwined movements: Department II prices go up and, thus, both departments workers' real wages go down. While the latter movement boils down to a forced savings (out of an income transfer from all workers' shrinking mass of wages toward department II expanding capitalists' net surplus), the former raise the rate of profit in department II and attract investments (out of that very expanding surplus) for it.

So in this case it is department II the one that casts a bottleneck on the system. In the short run it creates inflation and difficulties for the economy to keep growing; in the mid run, though, such (forced savings led) surplus is the source for financing the new investments that enable department II supply to catch up with its increased demand (Kalecki, 1954). Many underdeveloped countries that tried to implement a planned industrialization (via imports substitution for instance) had to cope with this type of unbalancing and rebalancing problem that among other things led to inflation and stagflation trends from time to time.

One could also argue that in an already matured and well-developed economy, not necessarily the heavy machines, infra-structure industries will be the only candidates for having low ceilings. Depending on the circumstances any one of the basic industries – not exclusively Erlich's sector Ia, or Kalecki's sector II – might occasionally be unable to increase its supply and cast a bottleneck that bogs down the whole economy.

In an economy where industries share the same wage ratio (the same "rate of exploitation") a higher capital-output ratio industry will realize a lower rate of profit. That means among all basic goods sectors, those with higher dead-to-living labor ratio would be the ones with a lower (value and -- indirectly -- physical) ceiling -- and thus they are prone to create bottlenecks.

We will see next that this idea is also familiar to a Sraffian system of commodities, in which industries with low physical net surplus are the ones with lower ceiling and thus the ones which regulate the short run (before adjustment) limit to growth. They and the adjustment they experience are then of particular interest for inflation matters. The Sraffian system bears also a preemptive advantage: it allows us to generalize the possibility of bottlenecks being casted by any basic good commodity so that we can investigate any of those possibilities.

6. Bottlenecks in a Sraffian Framework

In order to better formalize the foregoing disequilibrium/bottlenecks dynamics and assess its impact on inflation, it might be both insightful and useful to refer to a Sraffian-Pasinetti system of commodity production. Such system, unlike neoclassical traverse analysis, does not assume balanced growth, except as a benchmark.

To review in a glimpse its general framework, consider an all-basic goods economy in a certain period of time in which each *n* commodities are produced from labor and inputs of those *n* commodities themselves[8]. Let the input-output coefficients be the entries in the *n*x*n* Leontief input-output matrix **A** and the direct labor coefficients be the elements of the *n*-element row vector **L**. The quantities of commodity gross output, Q_1, Q_2, …, Q_i, …, Q_n, can be produced in many different proportions.

Each possible output combination (proportions) thus is to be understood in this context as a society's "choice". Those output quantities are the entries of an n-element column vector, **Q** so that the entries of $\mathbf{A^TQ}$ represent the physical quantities of capital goods needed to produce each of those commodity outputs – where the superscript "T" stands for the matrix transpose. In this sense the j^{th} entry of A^TQ is the quantity of commodity "j" that would be used up as inputs for producing all the other commodities, as well as for producing itself. The quantities net product -- surplus or net national income -- is therefore $\mathbf{Y} = (\mathbf{I}\text{-}\mathbf{A^T})\mathbf{Q}$. That means each industry generates a particular surplus, Y_i and thus realize its own rate of physical surplus, R_i. The j^{th} industry's R_j is equal to j's quantity net product divided by the quantity of "j" that has been used up as input by all industries so that, $R_i = Y_i / [Q_i - Y_i]$.

For the system to grow, and assuming an all-basic-goods economy, each industry will have to produce a surplus out of which a certain amount has to be channeled as new (additional) inputs – investments -- rather than consumed as final goods. If one of the

[8] For simplification's sake, at this level of abstraction, we assume all commodities to be basic ones and therefore being used up as inputs by the others as well as by itself. Basic goods in this Sraffian sense are raw materials, plants, equipments and wage goods. Non-basic goods would be luxury goods. Since the latter do not enter as inputs, they can expand production as long as basic goods (their inputs) can expand, which means their limit to grow is determined by basic goods. Therefore one can relax the all-basic goods assumption later without changing the conclusions we arrive here.

industries has not produced net surplus enough (if it has barely added up to the national income) it will create not only a bottleneck for itself but also for the other industries that are growing and using thus more of its commodity as essential means of production. That will hinder their and the whole economy's growth. On the same token, as Pasinetti would put it, "the minimum rate of surplus represents the growth bottleneck of the system: it represents the maximum rate at which the economic system can grow [steadily], given the proportions which have been chosen" (Pasinetti,1977, p.212).

Accordingly, "it is clear that the higher the rate of growth (…), the higher must be the minimum rate of surplus" (Pasinetti, 1977, p. 211). Since one cannot expect the technical conditions – i.e. inputs coefficients represented by entries in **A** and **L** -- to change overnight, the only way to reach a higher growth rate in the short-to-mid run is by changing the proportions within the output set, **Q**, so that the industry with the minimum R_i will produce more output -- and therefore more surplus out of it – at the expenses of another industry's surplus. It follows that "the range of possible choices open to the economic system is progressively narrowed down as we increase the rate of growth toward the maximum rate [because] an increase in the minimum rate of surplus necessarily requires a decrease in at least one of the other rates of surplus (…). The eventual outcome is clear (⋯) [i]n the limit," when there is no more "choices" but the very last one, the growth rate of the system and the aggregate surplus have reached their maximum, for all industries will have already relocated their surpluses offsetting one another until all rates of surplus "become equal to one another". Such final outcome is the Sraffa's standard system where, "no possibility of choice is open any more. The proportions become rigidly determined by the technology of the economic system" (Pasinetti, 1977, p.211).

The important point in such Pasinetti-Sraffa system is that the closer it is to its maximum rate of growth, the narrower the range of possible output proportions, and therefore the less room for sustainable variation in sectoral output proportions and growth rates. It follows that as the average growth rate rises toward the maximum rate, variations in actual output proportions and growth rates will face with increasing rigidity of the required proportions and growth rate. Bottlenecks become more and more frequent insofar as the economy average growth rate approaches the maximum growth rate.

In other words, the supply resistance becomes increasingly higher as the economy average throughput coefficient approaches one -- since the throughput is the actual-to-maximum growth ratio. Therefore the supply-resistance, rather than a linear function of the average throughput, must be a convex one: as such average goes up the function becomes increasingly positive -- i.e. its second derivative is positive. This is the same conclusion taken from the previously sketched inflation stylized facts as presented in section 5.

7. Throughput dispersion

A possible way to mathematically represent the abovementioned non-linearity story is to approximate it to a model that takes into account the across industry throughput dispersion. Beyond the average throughput, the dispersion of individual industrial sectors with regard to the mean may also reveal evidences of outliers. If, besides a high average throughput, much (or most) of those outlier (sector) throughputs lie to the right of the mean (average) in the frequency distribution, that must be a strong clue to the existence of bottlenecks already "narrowing the choices down" (cancelling out possible relieve rearrangements) and thus – "in the margin" -- increasingly boosting the supply resistance. Therefore, such resistance can be more accurately sensed if one includes to its function variables representing not only the first central moment, but also other higher moments such as the variance, the skiwness, the kurtosis and so on.

To do so let τ be the economy average throughput, $\tau = E[\tau_i] = \Sigma\tau_i/n$, where τ_i is the throughput of the i^{th} industry. The national, economy-wide, throughput is therefore the average growth strain of the several sectors, 1, 2, …, i , …, n, within the country's industrial chain.

To make sure bottlenecks are really bound to hold down the whole economy activity one should inspect the presence of outlier sectors with throughput much higher than the average ($\tau_i >> \tau$). A robust presence of such outliers might imply a high τ. But what to say about the σ_τ^2? A high variance may have an ambiguous meaning[9]. While it reveals outliers, by itself, it says nothing about which side of the distribution they lie on; besides, such variance may also signify room for proportions rearrangements (in a Pasinetti sense)

[9] Notice that there is no reason to assume any particular distribution in the across industry frequency distribution. It needs not be normal nor any other pre-determined type of frequency shape.

which may even relive rather than stifle the outlier bottlenecks[10]. Therefore, the sign and intensity to be assigned to the variance in the supply resistance function would depend on the higher central moments, particularly on the skewness and the kurtosis[11]. And that can be expressed by a multiplicative relationship between the former and the two latter moments. Accordingly, the supply resistance function of throughput (moments) is to be formalized as follows:

$$F[\tau, \sigma_\tau^2, ...] = \text{mean} + \text{variance [skewness} + \delta \text{ kurtosis} + (\dots)] \quad (8)$$

To make it short, one may only use the four first moments:

$$F[\tau, \sigma_\tau^2, \dots] = \tau + \sigma_\tau^2 \left[\left(\frac{E[\tau_i^3] - 3\tau\sigma_\tau^2 - \tau^3}{\sigma_\tau^3} \right) + \left(\frac{E[\tau_i^4] - 4E[\tau_i^3]\tau + 6E[\tau_i^2]\tau^2 - 3\tau^3}{\sigma_\tau^4} \right) \right] \quad (9)$$

Since it is not always easy to find and tackle data for each sector throughput in a country – and therefore to construct a dataset for each of its many moments -- it would be wise to simplify equation (8) by presenting it as a function of the average throughput only, F[τ]. Hence, equation (9) can be rewritten by expanding it,

$$F[\tau] = \tau + \frac{E[\tau_i^3]}{\sigma_\tau} - 3\sigma_\tau\tau - \frac{\tau^3}{\sigma_\tau} + \frac{E[\tau_i^4]}{\sigma_\tau^2} - \frac{4E[\tau_i^3]}{\sigma_\tau^2}\tau + \frac{6E[\tau_i^2]}{\sigma_\tau^2}\tau^2 - \frac{3\tau^3}{\sigma_\tau^2}$$

rearranging it

$$F[\tau] = \left(1 - 3\sigma_\tau - \frac{4E[\tau_i^3]}{\sigma_\tau^2}\right)\tau + \left(\frac{6E[\tau_i^2]}{\sigma_\tau^2}\right)\tau^2 + \left(\frac{E[\tau_i^3]}{\sigma_\tau} - \frac{\tau^3}{\sigma_\tau} + \frac{E[\tau_i^4]}{\sigma_\tau^2} - \frac{3\tau^3}{\sigma_\tau^2}\right) \quad (10)$$

and substituting the terms in the parenthesis by the coefficients ϕ_1 , ϕ_2 and ϕ_2, where

$-\phi_1 = \left(1 - 3\sigma_\tau - \frac{4E[\tau_i^3]}{\sigma_\tau^2}\right) < 0$, so that $0 < \phi_1$.

$\phi_2 = \left(\frac{6E[\tau_i^2]}{\sigma_\tau^2}\right) > 0$,

$\phi_3 = \left(\frac{E[\tau_i^3]}{\sigma_\tau} - \frac{\tau^3}{\sigma_\tau} + \frac{E[\tau_i^4]}{\sigma_\tau^2} - \frac{3\tau^3}{\sigma_\tau^2}\right) \cong 0$.

[10] Although such rearrangement, it takes time (in an Erlich sense) enough for one to be cautious in considering that relieve for inflation purposes

[11] Recall that the skewness is the degree of departure from symmetry of a distribution. Thus the more positively skewed is the distribution the more its "tail" is pulled in the positive direction. And the higher the Kurtosis, the fatter such tail would be.

The supply resistance function becomes then

$$F[\tau] = -\phi_1\tau + \phi_2\tau^2 \qquad (11)$$

and the inflation function in which the interaction between demand-pull and supply-resistance is expressed takes the form of:

$$\pi[e,\tau] = \theta_0 + \phi_0 F[\tau] = \theta + \phi_0 e(-\phi_1\tau + \phi_2\tau^2)$$

or simply put,

$$\pi[e,\tau] = \theta_0 - \theta_1\tau e + \theta_2\tau^2 e \qquad (12)$$

where the coefficients θ_0, θ_1 and θ_2, are all greater than zero so that

$\theta_0 = \theta$; $\qquad \theta_1 = \phi_0\phi_1$; $\qquad \theta_2 = \phi_0\phi_2$

Equation (12), a second-degree polynomial, shows us that the inflation curve might be represented as those depicted in **Figure 1**.

8. The Basic Model

Following all the reasoning detailed above, a general classical inflation model would look like as follows.

$$\pi = \gamma_i \pi_{-i} + \theta_0 - \theta_1\tau e + \theta_2\tau^2 e \qquad (13)$$

where π, θ_0, $[\gamma_i \pi_{-i}]$, τ, and "e" stand respectively for inflation, constant term, momentum (inflationary memory), throughput (private investment/profit) and new credit to the commodity market/GDP (Shaikh, 2001). Equation (13) shows the way the supply resistance, which is a function of τ -- as depicted in (12) -- amplifies (or stifles) the transformation of credit, "e", into inflation[12].

The multiplication of "e" by $F(\tau)$ allows us to capture the interaction between these two variables, which is a necessary condition for inflation. It seems clear that if there is no injection of new credit to the commodity market, there will not be pressures over prices

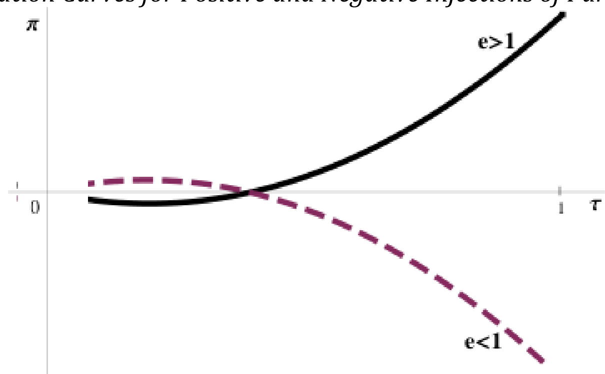
[12] Dummy variables can also be added in order to capture exogenous shocks. Lag variables should also be allowed as far as inflationary memory, hysteresis and other time factors are concerned -- let alone the delays between balanced growth rates and their impact on inflation.

no matter how high is the throughput.

9. Credit Outflow

All of that is valid whether "e" is positive or if it is negative. In case of a contraction of credit, e < 0, eF(τ) becomes an inverted-U, which can be seen in **Figure 1**.

Figure 1: *Inflation Curves for Positive and Negative Injections of Purchasing Power*



A high throughput might create deflation if an economy is facing with a severe credit crunch (negative injections of purchasing power). In this sense the adjustment scheme becomes upside down. To better understand such apparent paradox one should recall that neither throughput nor credit injection creates inflation (or deflation) by itself. It is only the concomitant presence of both that does so. Inflation happens when both have the same sign so that $e.F(\tau) > 0$; and deflation, when they have opposite signs, $e.F(\tau) < 0$.

10. Final Remarks

We have presented here an alternate approach to understand inflation based on the classical political economy tradition. It takes into account the interaction of excess demand and capital expansion hindrances for explaining upswing pressures over general prices needing not assume either neutrality of money, or full employment, or steady-state, balanced, equilibrium.

Only part of the new supply of money will really go to the commodity market (to the sphere of circulation) and count as a potential pressure over prices. But even that part of "active" money, which goes to circulation, will not necessarily put pressure over price by increasing aggregate demand over supply. We have unfolded two main points in this paper: (i) Money is endogenous and not only used for buying commodities but also for investing and expanding their production. An increase in Money supply has therefore a dual effect. If it generates both more demand and more supply of commodities roughly in the same proportion it will keep eventually prices unchanged. And (ii) the limit to grow is measured as the level of capital accumulation tightness, given the current rate of profit. Since capital accumulation is a rather uneven process the impact it (its interaction with injections of purchasing power) causes on inflation is a non-linear one.

For this reason we have delved further into the mechanisms of disequilibrium in a multi-sector system. The study of bottlenecks has shown pivotal for understanding such nonlinearity and thus for setting up a more appropriate functional form for an inflation model of this type.

Thus the throughput can either amplify or stifle the transformation of those injections into excess demand and consequently into inflation. In certain extreme cases it could still transform an inflow of credit into deflation.

Bibliography

Amendola, M. (1984) "*Toward a Dynamic Analysis of the 'Traverse'*". Eastern Economic Journal, Vol. 10, N02

Erlich, Alexander. (1967). "Notes on the Marxian Model of Capital Accumulation." American Economic Review, 57(2), pp. 599 615.

Filho, A.S. (2000), "*Inflation theory: A critical literature review and a new research agenda*". Emerald Group Publishing Limited, Vol. Iss: 18, pp.335 - 362

__________ (2002), "The Supply of Credit Money and Capital Accumulation : A Critical View of Post Keynesian Analysis In: Research in Political Economy. ed Paul Zarembka, pp. 309-334.

Handfas, A. (2012)."Demand Pull Supply Resistance in a Classical Inflation Model" New School for Social Research Dissertations: http://gradworks.umi.com/35/11/3511238.html

Kalecki, M. (1954) "The Problem of Financing Economic Development". in J. Osiatynsky, ed., Collected Works of Michal Kalecki, Vol. V Oxford: Oxford University Press, 1993.

__________ (1966) "The Difference Between Crucial Economic Problems of Developed and Underdeveloped Non-socialist Economies". Em J. Osiatynsky, ed., Collected Works of Michal Kalecki, Vol. V Oxford: Oxford University Press, 1993.

Keynes, J.M. (1937) "The General Theory of Employment", The Quarterly Journal of Economics, Vol. 51, No. 2

Marx, K. (1976). "Capital – Volume I. London: Penguin Classics, 1992.

Pasinetti, L. "Lectures on the Theory of Production". Columbia University Press, 1977

_______ (1978). "Capital – Volume II. London: Penguin Classics, 1992.

Shaikh, A., (2014) (forthcoming book).

________ (2011). "The Empirical Linearity of Sraffa's Critical Output-Capital Ratios" in Festschrift volume in honor of Heins Kurz .

_________ (1999), "Explaining Inflation and Unemployment: An Alternative to Neoliberal Economic Theory", in Contemporary Economic Theory, Andriana Vachlou (ed.), Macmillan, London. 1999.

Sraffa, P. (1960) "Production of Commodities by Means of Commodities - Prelude to a Critique of Economic Theory", Cambridge, Cambridge University Press.