# Causes and Effects of Monetary Disequilibrium in Ricardo and Thornton

By

David Glasner

Federal Trade Commission

March 2011

# Causes and Effects of Monetary Disequilibrium in Ricardo and Thornton

## I Introduction

It is an honor to have been invited to participate in this conference on the monetary theory of David Ricardo. Having left the academy many years ago to become a government economist, I am grateful to be here in the company of so many distinguished Ricardian scholars, whose knowledge of and expertise in the vast output of the great man whose contributions to economic science we seek to perpetuate, undoubtedly far exceeds my own.

That I find myself in your company and am addressing you today is presumably the result of four papers on classical monetary theory (Glasner 1985, 1989, 1992, 2000) that I wrote in the previous century, in two of which I explicitly discussed Ricardo's monetary theory, but only to show that his views on monetary questions were more or less in accord with a general monetary paradigm that I had designated "classical monetary theory," to distinguish it from the more familiar quantity theory of money, which has, without appropriate qualification, routinely been ascribed to the classical economists.

In brief, I argued that the characteristic feature of the classical monetary theory was a privately produced, unregulated, supply of money convertible into a real asset, usually gold.[1] In contrast to the standard version of the quantity theory, in which the quantity of money is exogenously determined (at least in the short run), I argued that the quantity of money in the classical theory was endogenous. Not only was it endogenous, but there were two mechanisms

[1] Ricardo, in fact, favored limiting the convertibility of notes to gold bullion, which, except for a small class of international transactions, did not serve as a medium of exchange, only a standard of value. So Ricardo's objective was clearly to minimize the monetary demand for gold, and to prevent a demand for gold for monetary purposes from having a destabilizing, i.e., deflationary influence on the economy.

operating independently to ensure endogeneity. First, the balance of payments ensured that an excess supply of money would cause the export of specie abroad, but, additionally, if any bank issued its banknotes to excess or created too many deposits, it would suffer a loss of assets and profits as a result of a negative balance at the clearinghouse.[2] Quantity theorists in the Humean tradition would argue that the endogeneity ensured by the balance of payments was only in the medium to long-term, short-term overissues of banknotes being a possibility of greater or lesser import, depending on the circumstances and the theoretical perspective of the analyst. However, in the classical monetary tradition whose existence I attempted to document and to which, I believe, Ricardo cleary belonged, even a temporary overissue was ruled out by the integration of world markets, prices of internationally traded articles always tightly constrained by international competition. Excess supplies of money (or, more precisely, of money-backing) would add to the local demand for tradable goods, causing a negative trade balance until the excess supply was eliminated without affecting domestic prices (at least for tradables).[3]

That the Bank of England had a special privileged position in the monetary framework of Great Britain might (or might not) be rationalized within such a theory, but to do so involved no rejection of the basic paradigm of an endogenous money supply and an exogenous price level (anchored by convertibility and international commodity arbitrage), nor did it render the quantity theory the appropriate theoretical model for analyzing the British monetary system under

[2] Given the latter mechanism, gold flows reflected changes in the relative demands for gold for non-monetary uses or to be held by the banking system as reserves or for coinage, not excess demands by the public for money to hold. The equilibration of the supply of with the demand for money to hold was achieved by a domestic monetary mechanism. As I have observed elsewhere (Glasner 1985, Glasner 1992) this mechanism has bearing on our understanding of how Say's Law could obtain even in a monetary economy.

[3] While the prices of tradables are tightly constrained by the forces of international commodity arbitrage to nearly continuous equality across space, non-tradables are constrained only indirectly by the tendency for profits to be equalized in all industries in all countries. That equalization may occur only in the intermediate or long run.

convertibility. The analysis becomes more complicated under inconvertibility, but that complication is irrelevant for purposes of this quick overview.

I would note in passing that the motivation for my explorations into classical monetary theory was less an interest in the history of economic thought in general or the history of monetary thought in particular than my realization, as an undergraduate taught by Ben Klein, that traditional arguments for why competition in the supply of money is unworkable were logically defective. It was that conceptual insight coupled with the explicit suggestion of another one of my teachers, the late, great, and sorely missed, Earl Thompson, that led me to explore the work of the classical monetary economists to see whether they had had any awareness of the conditions under which a privately produced competitive supply of money was workable, notwithstanding the subsequent widespread acceptance of unfounded claims that free competition would drive the value of money down to zero.

Among the critical comments elicited by my earlier papers, a recurring theme was that I was trying to force the ideas of the classical economists into a theoretical framework that they would not have recognized, much less accepted. I have particularly in mind the comments of two distinguished economists and historians of thought, whose contributions I greatly admire, Mark Blaug and D. P. O'Brien. In one of the papers referenced above (Glasner 2000) I attempted to respond to their specific criticisms, and do not propose to reopen that discussion here. But I cannot help making a general methodological point: to do history of thought (or the history of science generally) correctly, it helps to have a correct theory (or rather a more nearly correct theory) than the one employed by the economists or scientists under historical consideration. If we can see further than our predecessors, it is only because we stand on their shoulders. But just because we stand on their

shoulders and see what they could not, we can also, if we make the effort, gain insights into the logic of their theories and explanations that they themselves did not necessarily have.

Rather than continue in this perhaps overly philosophical vein, let me now briefly outline the topic that I wish to discuss today. While my earlier papers were mostly concerned with working out the classical theory of the supply of money produced by a competitive banking system, I think that it is also worth considering the classical theory of the demand for money, and particularly the theory of the demand for money in Ricardo and Thornton. The classical theory of the money supply, though only partially developed, seems to me an important contribution inadequately transmitted to and developed by the successors of the great classical monetary theorists. The classical theory of the demand for money, despite perceptive observations by individual theorists, notably by both Henry Thornton and David Ricardo, offering sophisticated insights into the reasons for holding money in general and for holding some monetary instruments rather than others, seems to me to have been fragmentary and unsystematic. Lacking a clearly developed and systematic theory of the demand to hold money as an asset, a theory that only gradually took shape in the 1920s and 1930s, before Keynes made it the cornerstone of his *General Theory*, Ricardo, in attempting to demonstrate that only an excess issue by the Bank of England could have caused a depreciation of sterling relative to bullion, could not formulate the theoretical problem that he was trying to solve in terms of the supply-demand framework that would have been necessary for him to arrive at the solution to his problem. Similarly, Thornton's remarkable natural rate theory, prefiguring the subsequent theories of Wicksell, Marshall and their many offspring, but lacking a theory of the demand for money, was incomplete at a crucial step in the argument: the demand to hold money is missing from his model. Perhaps aware of the logical shortcomings of the other's theories, Ricardo and Thornton seem never to have engaged each other directly on the points of difference in their seminal

contributions. How each of these two giants of economic theory viewed each other's contributions seems destined to remain one of the tantalizing unanswered questions in the history of economic thought.

## II Ricardo on Depreciation and Overissue

David Ricardo's first contributions to economic theory were made in his letters to the *Morning Chronicle* in 1809, later revised and published as *The High Price of Bullion* (1810). As a fervent advocate of the gold standard, Ricardo, having missed the first round of the Bullionist

Debates after the Bank of England suspended, with the approval of the government, convertibility of its banknotes into gold, reopened the debates by arguing forcefully in his letters to the *Morning Chronicle* that the recent increase in the price of bullion had been caused by an overissue of Bank of England notes.

In attacking the policy of the Bank of England, Ricardo made an exceptionally strong claim. Not only had the rise in the price of bullion in 1809 been caused by an overissue of notes by the Bank of England, a charge previously advanced by others, notably Walter Boyd (18010, in the first round of debates in 1800, on which Henry Thornton had expounded at length in his great work of 1802, Ricardo went much further, contending that, as a matter of strict economic logic, an increase in the sterling price of bullion could not have been caused by anything but an increase in the quantity of Bank of England notes beyond that amount which could have circulated under convertibility. This proposition had not been advanced at all in the first round, and only John Wheatley (1807), writing during the period of relative calm between the first and second rounds, had ever argued for the proposition before Ricardo did in 1809.

Thus, even those who had held the Bank of England responsible for the depreciation of sterling (again with the exception of Wheatley) had not argued that sterling could never depreciate relative to gold independently of an increased issue of Bank of England notes. Two contingencies were mentioned as possible exceptions to the presumption that a depreciation of sterling was caused by an overissue of banknotes: 1) a poor harvest requiring grain imports to replace domestic production, or 2) remittances to continental allies that would join the war against Napoleon.

Because gold would have to be sent abroad to finance the payments for grain or for military supplies for continental allies, inconvertible banknotes being generally unacceptable outside of Britain, the necessary gold with which to effect the desired transactions could be obtained only by bidding up the price of gold in terms of sterling. This reasoning was sufficiently persuasive for Thornton to have accepted it, as did the first important Bullionist writer Walter Boyd, and to have been endorsed by the Parliamentary Report recommending the quick restoration of convertibility, the famous Bullion Report, co-authored by Thornton himself with Horner and Huskisson. Ricardo, however, rejected it completely. Let us try to understand Ricardo's reasoning.

Ricardo's theoretical approach, as we know, was to work out a barter analysis, abstracting from money, and then to use the barter analysis as a benchmark with which his monetary analysis had to conform. In the barter theory of international trade, it is exports that constitute the demand for imports. Thus, if Britain had to import grain because of a domestic crop failure, the ultimate means of payment for the grain imports would be Britain's exports of other products. On the one hand, this seems to ignore the reality that a crop failure might occasion extraordinary purchases of grain that could not immediately be paid for by increased exports, but would have to be paid for in first instance by gold shipments. But Ricardo held that gold shipments would not be resorted to immediately. Rather anticipating that British exports would increase to finance grain imports,

foreign lenders would lend short term to cover the immediate demand for grain imports. The spot exchange rate for sterling might fall to the gold export point, but an expectation that future British exports would increase to finance the exceptional imports would imply a forward exchange rate above the spot rate, making borrowing gold or foreign currency today for future repayment cheaper than the costs of shipping gold. Meanwhile, because prices of internationally tradable products were constrained by arbitrage from deviating from each other across countries by more than transportation costs, falling home demand for non-grain products to finance increased grain imports would cause their prices in Britain to fall to the point at which either exports from Britain would become profitable or imports to Britain would become unprofitable, thereby generating the export surplus expected by the market and require to finance the added imports of grain even without the export of specie.

Let me observe in passing that Ricardo's reliance on the tight integration of world markets for internationally traded goods shows clearly that Ricardo reasoned in terms of a common international price level under a gold standard. "The exportation of the coin," wrote Ricardo (1810, 61) "is caused by its cheapness, and is not the effect, but the cause of an unfavourable balance: we should not export it, if we did not send it to a better market, or if we had any commodity which we could export more profitably." Commenting on Ricardo's rejection of "Thornton's assertion that an unfavorable balance of payments could cause an outflow of gold on the ground that it would first have to be explained why other countries were unwilling to accept the deficit in commodities rather than exclusively in gold," Hayek ([1929] 1991, p. 200] defended Thornton because "the only way payment for the extra imports can lead to an increase in exports is through the outflow of gold, which eventually produces a lowering of the domestic commodity prices."

Clearly adopting the view that adjustment to international monetary equilibrium can be achieved only by way of the Humean price-specie flow mechanism in which adjustments in local price levels induce the required changes in commodity flows, Hayek here failed to see that Ricardo conceived of an adjustment mechanism involving no change in local price levels and achieved entirely by commodity flows driven by arbitrage opportunities within limits determined by the transport costs of individual commodities. Ricardo, in his *Reply to Bosanquet*, had noted how the efficient functioning of the foreign exchange and gold markets, constantly eliminating profit opportunities by buying in cheap markets and selling in dear ones, equalized the value of gold across all international centers. Considering commodity markets to be equally as efficient as the gold market in arbitraging local price differences for the same commodity in different locations, Ricardo, therefore, believed that any slight decrease in commodity prices in Britain owing to the reduced British demand for commodities in general occasioned by a bad harvest or foreign remittances, would cause British exports to increase. British purchases of grain abroad would necessarily create profitable opportunities for the export of British commodities even before it became profitable to ship gold.

The logic of Ricardo's argument that a bad harvest or foreign remittances could not cause a depreciation of sterling in terms of gold seems to me unassailable as far as it goes. And yet, I do not think that Ricardo succeeded in attempting to refute the proposition, a proposition accepted by Thornton and by most of the Bullionist writers, that a bad harvest or foreign remittances could cause a depreciation of sterling independently of any overissue by the Bank of England. What is lacking in Ricardo's argument?

The problem with Ricardo, as I suggested above, is a failure to tease out the implications of a bad harvest or of foreign remittances for the demand for money. We now take for granted that the

demand for money to hold as an asset is a function of many variables, which would certainly include, among other items, the level of income and wealth. Thus, insofar as a bad harvest or foreign remittances reduced the aggregate income and wealth of Great Britain, as they surely would have, a bad or harvest or foreign remittances would very likely have implied a reduction in the demand to hold money, which, for our purposes, we may identify with the demand for inconvertible Bank of England notes. It is an elementary proposition of monetary theory that a reduction in the demand for money with an unchanged supply implies an excess supply of money, an increased price level, and an increased price of gold, i.e., a high (relative to the pre-restriction par value) price of bullion. That Ricardo failed to see this straightforward, and, from our perspective, even obvious proposition, shows that Ricardo's analytical apparatus had not adequately assimilated the notion of a demand for money to hold as an asset as part of a general analysis of the level of prices and the exchange rate.

Now one might defend Ricardo against the charge that he failed to recognize that the disturbances he was analyzing implied a reduced demand for money by positing that Ricardo must have regarded a failure by the Bank of England to reduce its note issue, in the event of a reduced demand for Bank of England notes, to be equivalent to an increase in the note issue of the Bank of England. While such a response is certainly possible -- and I leave it to others to decide whether it would be consistent with Ricardo's assertions that only an *overissue* of notes by the Bank of England could cause a depreciation of sterling relative to gold, I think that once the discussion about Ricardo's position has reached this point, the discussion has crossed the line from substance to semantics, and we are no longer assessing an argument that might be valid or invalid but .

While on the subject of a falling demand for money as a possible cause of depreciation, I would observe that there was an entirely different reason why the demand for Bank of England

notes might have fallen after the Restriction, which was that the Restriction could itself have engendered expectations that the future value of Bank of England notes would fall after the link between the value of the notes and the value of gold had been severed. In his classic treatment of the Bullion Debates, Jacob Viner (1937, 132) noted with some degree of surprise the absence of any such argument by the Bullionist critics of the suspension of convertibility. Viner explained the absence of this argument from the Bullionist criticisms of the suspension by suggesting that to have made such an argument would have been politically dangerous in the charged atmosphere of the time when England was engaged in military conflict with the French. To have called into question, even indirectly, the soundness of sterling by attributing a loss of its value to a lack of confidence in its future value would, Viner conjectured, have left the Bullionists open to a charge of that they were bringing the national currency into "discredit" when England was at war. The argument seems plausible on its face, but Viner, himself, acknowledged that it was only conjecture. However, one cannot exclude the possibility that the Bullionists did not suggest that sterling might have depreciated against gold because of expectations of future inflation as a result of the suspension of convertibility because their conception of the demand for money was not sophisticated enough to have accommodated this idea within their analytical tool kit.

## III Thornton on the Natural Rate of Interest and Inflation

Perhaps it is only fair that, after a century of undeserved obscurity, Henry Thornton now stands out as the premier monetary theorist of the classical period.[4] I have no interest in arguing about who was the greatest classical monetary theorist, and I have no quarrel with anyone who

[4] Skaggs has argued persuasively that Thornton's obscurity for most of the nineteenth and the early part of the twentieth century was exaggerated and that Thornton actually was neither forgotten by nor without influence on the likes of Mill and Bagehot and others who helped formulate what constitute British monetary orthodoxy at the end of the nineteenth and early twentieth century.

would award the laurel to Henry Thornton, especially inasmuch as he seems to have been an exceptionally admirable human being. However, I do propose to argue in this section that what is generally considered to have been Thornton's greatest contribution, his essentially full anticipation of Wicksell's natural rate model almost a century before Wicksell himself expounded it himself, was logically incomplete and possibly invalid (as indeed was Wicksell's model). Here again, the sources of the problem turn out to have been an inadequately developed theory of the demand for money and lack of attention to international commodity arbitrage. After explaining what I think is wrong with the Thornton/Wicksell natural rate model, I will offer some thoughts on why Ricardo seems not to have been persuaded by Thornton's model although he was surely well acquainted with it.

The natural rate model is well enough known for a very brief verbal summary of the model to suffice as an introduction to the discussion. At any moment, there is a range of potential investment opportunities open to entrepreneurs. The minimum rate that entrepreneurs regard as sufficient to warrant undertaking any of the prospective investment projects under consideration may be described as what Wicksell called the natural rate of interest or what Thornton called the rate of profit, or what Keynes called the marginal efficiency of capital. If the banking system offers to lend at a rate below the rate of profit or the natural rate or interest, entrepreneurs will seek to borrow funds from banks on favorable terms, and, as a result, the banks will create increasing amounts of money (either notes or deposits) in lending to eager borrowers. The additional money created will force up prices, and if interest rates are not raised, the process will continue without limit.[5]

[5] Actually, although Thornton held that there was no limit to the amount a bank could lend if it offered to do so a lending rate below the natural rate, that conclusion presumes that there unlimited seemingly profitable investment projects available at the interest rate at which banks are prepared to lend. But there is nothing that guarantees that there the number of such projects available is unlimited. The number of such projects may in fact be finite, and even if the number is, for

Thomas Humphrey (1997), relying on David Laidler's (1972, [1975]) encapsulation of the Wicksellian model, summarized the Thornton/Wicksell model in a system of five equations. Borrowing Humphrey's notation, I write the system in terms of the following variables, real investment, $I$, savings, $S$, the rate of interest charged by banks on loans, $i$, the natural rate of interest (the profit rate on investment), $r$, excess aggregate demand, $E$, the change in the money stock, $dM/dT$, and the change in the price level, $dP/dT$. The five equations describing the dynamics of the system are the following

$$I - S = a(r - i),\ a > 0;$$

$$I - S \equiv dM/dT$$

$$I - S \equiv E$$

$$dP/dT = kE,\ k > 0$$

$$di/dT = b(dP/dT),\ b > 0$$

Equation (1) says that investment exceeds saving insofar as the natural rate exceeds the banks' lending rate. Equation (2) says that funds that banks advance to businesses to finance their investment projects beyond the amount of private savings represent newly created money, so that the excess of investment over savings is equivalent to the increase in the money supply. Equation (3) says that the gap between investment and savings is identically equal to excess aggregate demand. Equation (4) says that the rate of inflation is proportional to excess demand. Finally equation (5) says that if banks set their lending rate below the natural rate, the pressure of excess

---

practical purposes, unlimited an attempt to realize them all simultaneously would drive up factor costs, thus reducing what Keynes called the marginal efficiency of investment below the marginal efficiency of capital.

demand eventually forces banks to raise the lending rate to match the natural rate, thereby restoring an equilibrium with a stable price level and no excess demand.

A few substitutions allow lead to the following relationships

$$dP/dT = ka(r - i)$$

and,

$$dM/dT = a(r - i).$$

Equations (6) and (7) relate the rate of monetary growth and the rate of inflation to the difference between the natural rate of interest and the lending rate charged by the banks. Because equation (5) tells us that the lending rate converges to the natural rate, the rate of monetary expansion and the rate of inflation also converge to zero, at which point the equality between savings and investment, and hence macroeconomic equilibrium, is restored.

What appears puzzling to me about this system is that while it contains a variable representing the supply of money, it has no variable for the demand to hold money and thus no equilibrium condition requiring that the amount of money demanded equal the amount of money supplied. The missing demand for money variable and the missing equilibrium condition were not left out owing to an oversight by Laidler and Humphrey. They were accurately representing the Thornton/Wicksell model in which the demand for money has no role to play, or, perhaps on an alternative interpretation, is implicitly assumed to adjust to whatever amount of money is supplied.

Perhaps one way to understand the system is that the real demand to hold money is a constant, implying that the increase in the price level following an increase in the quantity of

nominal money is the rise in prices necessary to keep the real quantity of money constant.[6] This would still leave the real quantity of money invariant with respect to the interest rate, a version of the demand for money not very congenial to modern sensibilities, but often attributed, mistakenly in my view, to classical monetary theorists. This difficulty could presumably be overcome by making $k$ in equation (6) a function of the rate (or rates) of interest. So, if the amount of money supplied were to increase owing to an increased natural rate, gradually raising (as specified in equation (5)) the market interest rate, the quantity of money demanded would fall owing to the increased interest rate. Because the quantity of money demanded falls, the increase in prices required to reestablish equilibrium between the demand for money and the supply would have to be more than proportional to the increase in the nominal quantity of money.

While it would be possible to add a demand for money equation and an equilibrium condition to the Thornton/Wicksell model in the way that I just outlined, doing so would still leave another, and I think deeper, problem with the model unresolved. The other problem that I have with the natural rate/market rate paradigm is that the cumulative process it describes -- banks increasing lending to businesses and entrepreneurs in response to an increase in the profit rate expected by entrepreneurs -- involves an essentially mechanical response by the banks to the increasing demand for bank loans, a response not grounded securely in the self-interest of the

[6] In his own unique variant of the Thornton/Wicksell model, R. G. Hawtrey derived the effect of a reduced bank rate from its effect on the incentives of traders to add to their stocks of goods in process and by doing so to increase consumers' incomes and outlays, not by creating an excess supply of money that would raise prices. "The intention [of the Bank of England], wrote Hawtrey (1938, p. 37) "was, by varying the discount rate, to regulate the amount of bills brought to the Bank of England, and so to control the volume of currency against which those bills were held. But the policy was by no means dependent on the quantity theory. In fact the discount rate modified the quantity of currency by modifying activity; the latter was cause and the former was effect. The link between the quantity of currency and the price level existed (though not in so rigid a form as was then supposed). But the variations in the quantity of money and the price level were effects of a common cause, the variations in the consumers' income, and it was these latter that were influenced by the discount rate."

engines of the process, the individual banks themselves. With no increase in the demand to hold money, increased lending by any single bank would cause interbank clearings to turn against that bank unless it simultaneously took steps to induce the public to increase their holdings of its liabilities. The most direct, though not the only, way to induce the public to hold additional liabilities of the bank would be to raise the interest paid on deposit balances held at the bank.

Now it is true that increased lending by other banks would tend to offset the adverse clearings against any single expanding bank occasioned by its own increase in lending, but there would still be net leakages out of the banking system, either internally via demand for coin or externally via a demand for bullion, so that parallel expansion by the whole banking system would not fully insulate individual banks against pressure on their reserve position when simultaneously increasing their liabilities with no equivalent increase in the amount of those liabilities demanded by the public. Banks would therefore feel pressure to raise their deposit rates to maintain or replenish their reserves, and, before long, to raise their lending rates as well, thereby limiting the increase in their lending as rates rose toward the expected entrepreneurial profit rate. In the Thornton/Wicksell model, the only pressure to increase the lending rate stems from the tendency for prices to rise in response to the monetary expansion fueled by the banks. But consideration of the incentives of the banking system suggests that the pressure to raise interest rates would be felt even before any price level effects were evident.[7]

[7] The discussion in this paragraph does not dispute that banks can expand to some degree if they do so in concert. This is certainly true, but it is a less compelling argument than conventional banking theory seems to take for granted. What is it about banking that allows one to assume with little or no supporting argument that an individual bank would expand conditional on the assumption that all other banks will do the same, irrespective of the number of other banks in the system? Would one casually assume that individual non-banking firms will raise their prices conditional on the assumption that every competing will also raise its price, irrespective of the number of other competitors in the market? What accounts for the difference in the conventional assumption about

Yet another problematic aspect of the Thornton/Wicksell model is the unspecified and ambiguous relation between domestic prices, external prices and the exchange rate. Such ambiguity was absent from Ricardo's discussion because he emphasized the constraint on domestic prices imposed under the gold standard by international commodity arbitrage, limiting the extent to which increased bank lending in any one country could raise the price level in that country. The inflationary effects in one country, as Adam Smith ([1776] 1976) had earlier observed about the substitution of bank notes for coin, are exported, and dissipated, via commodity arbitrage, and the effects of reserve drains on private banking positions are, via the processes mentioned in the preceding paragraph, likely to limit increases in lending, driving up both deposit and lending rates before any significant inflationary effects are perceptible.

One might also observe in this connection that Thornton's discussion was, at least in one important respect, superior to Wicksell's, namely, Thornton offered his analysis in the course of a discussion of an inconvertible currency with a binding constraint on bank lending rates, owing to the 5 percent usury limit on the interest rate that banks charged borrowers. Wicksell, on the other hand, explicitly argued that the irrespective of whether bank liabilities were directly or indirectly convertible into gold, movements in the price level were governed solely by the relation between the bank lending rates and the profit rate expected by entrepreneurs.

It is also noteworthy that Thornton, in expounding his version of the natural rate theory, was mainly concerned to show that simply following the real-bills doctrine as the directors of the Bank of England claimed to have been doing. He had explained earlier in his volume that rising prices and a premium on bullion might be caused either by an overissue by the Bank of England or

---

how competing firms and competing (?) banks make their decisions? I am unaware of any discussion of this issue in the literature on banking theory.

by a balance of payments deficit resulting from a bad harvest or overseas remittances. Defenders of the Bank of England had argued that if the Bank of England observed the real bills doctrine, an overissue would not be possible, so that if there were a depreciation of sterling and a rise in prices, it could only have resulted from an unfavorable balance of payments. It was to refute this anti-Bullionist argument that Thornton advanced his natural rate theory, showing that even if the Bank had restricted itself to lending on the security of real bills, the 5-percent ceiling on the interest the Bank could charge for loans would induce entrepreneurs to borrow in unlimited quantities from the Bank if the profit rate expected were greater than the 5 percent rate that the Bank of England could legally charge on loans. The point therefore was to show that although an unfavorable balance of payments might be the cause of a depreciation of sterling, observance of the real bills doctrine was no guarantee against overissue by the Bank of England, so that the assertion of the anti-Bullionists that a depreciation of sterling had to have been caused by an unfavorable balance of payments was invalid.

For Ricardo the whole exercise was unnecessary. Believing that he had demonstrated that a depreciation of sterling was impossible except as a result of overissue, he had no reason to refute the real bills doctrine. Refuting the real bills doctrine was only necessary if one conceded that a premium on bullion might be caused by something other than an overissue. Ricardo never made that concession. Moreover, Ricardo must have objected to the substance of Thornton's treatment. As Ricardo relied on international commodity arbitrage to obviate financing overseas purchases with bullion, he could not have agreed with Thornton that a bank lending rate pegged below the profit rate would lead to an increase in domestic prices, and then, as a result, an outflow of gold and a depreciation of sterling. As long as the pound was convertible into gold, international commodity arbitrage would keep British prices equal to those abroad in terms of gold equivalents. If the

convertibility of Bank of England notes were suspended, the depreciation of sterling would measure the extent of overissue. From Ricardo's perspective, the natural rate/market rate distinction was entirely otiose.

## IV Conclusion

I have explored in this paper some of the differences that separate two giants of early classical monetary theory and the reasons why they took different positions. A primary reason for their differences has turned out to be Ricardo's consistent adherence to the principle that international commodity arbitrage would equalize commodity prices across all countries operating on the gold standard. Despite Thornton's recognition of this principle at some points in his great work, he did not maintain that position consistently, lapsing into arguments that relied on Humean relative price level adjustments. This was especially true of Thornton's rendition of the natural rate/market rate theory in which the proposition that overissue could result in a depreciation of sterling was established by way of a Humean adjustment process that to Ricardo must have seemed fallacious.

While Ricardo may have had valid reasons for rejecting Thornton's most celebrated contribution to monetary theory, his own analysis of the possible causes of a depreciated inconvertible currency was not flawless either. That error, I suggested, would have been evident to Ricardo if he had been working with a more clearly defined concept of the demand for money to hold. It similarly appears that Thornton's exposition of his natural rate theory was also marred by a faulty, or at least incomplete, understanding of the demand for money to hold.

Before bringing this discussion to a close, I would offer one further observation on the relation between Ricardo's and Thornton's monetary theories. In Ricardo's account, the value of

money can be said to be determined either by the value of gold as a commodity, with a definite demand encompassing both monetary and non-monetary uses. Arbitrage between monetary and non-monetary uses equalizes the value of gold across all uses. Banks affect the value of money only insofar as they reduce the total demand for gold by allowing gold to be withdrawn from some monetary uses and to be devoted instead to non-monetary uses of lower value than was otherwise possible. Thornton's analysis is less clear cut than Ricardo's. By increasing the quantity of money in the short term, banks can raise prices even under a gold standard even without causing a reduction in the monetary demand for gold.

A similar analytical difference can be discerned today between those who argue that the value of money is determined essentially by the supply of and demand for high-powered money, which I define as non-interest-bearing currency and reserves, and those who argue that the value of money is determined by the supply of all kinds of money and near-moneys and the demand for all such instruments, and whether or not those instruments bear interest. The first approach might be called Ricardian and the second Thorntonian.

An interesting parallel between the Ricardian and Thorntonian approaches may be found in the more recent controversy (now, to my amazement, almost 50 years old) between what was then called the New View of banks and financial intermediaries and the Old View. Advocates of the Old View generally thought that they were defending the notion that monetary policy could affect aggregate demand and the price level by adjusting through changes in the central bank's discount rate to commercial banks or through open market operations while advocates of the New View argued that in a proper general equilibrium model there was no meaningful distinction between the monetary liabilities of banks and the non-monetary liabilities of other financial intermediaries. Advocates of the New View were therefore supposed to be arguing, and supposed themselves to be

arguing, that monetary policy did not matter. But if I am right in suggesting that the New View were really Ricardians (treating the financial sector including banks as irrelevant to the determination of the price level) and the Old View were Thorntonians (treating movements in the price level as the result of expansion or contraction by the banking system), then the crux of the argument was whether the effects of monetary policy can be analytically reduced to the supply and demand of high-powered money or what is now called the monetary base or whether some aggregate measure of all monetary instruments (of varying degrees of "moneyness") must be juxtaposed against some aggregate measure of the demand to hold such instruments. Whether monetary policy is effective or not is orthogonal to this issue. But at the time, everyone involved seemed to think that the dispute mattered only because of its policy implications.

Having disposed of the bogus policy issue, I think that a dispassionate assessment of the analytical arguments would show that the New View clearly had the better analytical grasp of the role of banks and financial intermediaries in the economy. The subsequent policy success of Monetarism (though short-lived) unfortunately cut short the promising analytical progress that the New View had started to make. But this is a story that will have to be told on some other occasion.

REFERENCES

Arnon, A. 2004. *Monetary Theory and Policy from Hume and Smith to Wicksell.* Cambridge: Cambridge Univ. Press.

Boyd, W. 1801. *A Letter to the Right Honourable William Pitt, on the Influence of the Stoppage of Issues in Specie at the Bank of England, on the Prices of Provisions, and other Commodities.* London.

Frenkel, J. and H. G. Johnson. 1976. "The Monetary Approach to the Balance of Payment: Essential Concepts and Historical Origins." In *The Monetary Approach to the Balance of Payments*, edited by J. A. Frenkel and H. G. Johnson, pp. 21-45. Toronto: Univ. of Toronto Press.

Glasner, D. 1985. "A Reinterpretation of Classical Monetary Theory. *Southern Economic Journal* 52:46-67.

Glasner, D. 1989. "On Some Classical Monetary Controversies." *History of Political Economy* 21:201-29.

Glasner, D. 1992. "The Real Bills Doctrine in the Light of the Law of Reflux." *History of Political Economy* 24:867-94.

Glasner, D. 2000. "Classical Monetary Theory and the Quantity Theory of Money." *History of Political Economy* 32:39-59.

Hawtrey, R. G. 1938. *A Century of Bank Rate.* London: Longmans.

Hayek, F. A. [1929] 1991. "The Period of Restrictions, 1797-1821." Chap. 11 in *The Trend of Economic Thinking. The Collected Works of F. A. Hayek.* edited by W. W. Bartley III and S. Kresge. Indianopolis, IN: The Liberty Fund.

Hollander, S. 1979. *The Economics of David Ricardo*. Toronto: Univ. of Toronto Press.

Humphrey, T. 1997. "Natural Rate of Interest." In *Business Cycles and Depressions: An Encyclopedia*, edited by D. Glasner, pp. 476-79. New York: Garland.

Laidler, D. E. W. [1972] 1975. "On Wicksell's Theory of Price Level Dynamics." Chap. 5 in *Essays on Money and Inflation*. Chicago: Univ. of Chicago Press.

Ricardo, D. [1809] 1951. *The Price of Gold: Three Contributions to the Morning Chronicle.* In *The Works and Correspondence of David Ricardo*, edited by P. Sraffa, pp. 13-46. Cambridge: Cambridge Univ. Press

Ricardo, D. 1810. *The High Price of Bullion*. In *The Works and Correspondence of David Ricardo*, edited by P. Sraffa, pp. 47-127. Cambridge: Cambridge Univ. Press

Ricardo, D. 1811. *Reply to Mr. Bosanquet.* In *The Works and Correspondence of David Ricardo*, edited by P. Sraffa, pp. 155-256. Cambridge: Cambridge Univ. Press

Skaggs, N. 1995. "Henry Thornton and the Development of Classical Monetary Economics." *Canadian Journal of Economics* 28:1212-27.

Thornton, H. [1802] 1939. *An Inquiry into the Nature and Effects of the Paper Credit of the Bank of England,* edited by F. A. v. Hayek. London.

Viner, J. 1937. *Studies in the Theory of International Trade*. New York.

Wheatley, J. 1807. *An Essay on the Theory of Money and the Principles of Commerce*, vol. I. London.