Are Trade Deficits a Problem?

K. Alec Chrystal and Geoffrey E. Wood

In 1986, the U.S. trade deficit exceeded $140 billion. Such substantial trade deficits often are considered a sign of weakness in the economy. While this situation is something of a novelty for the United States, many other countries have had trade deficits off and on throughout the postwar period.

The purpose of this article is to explain what is meant by trade deficits within the context of the balance of payments, to outline the circumstances under which the state of the balance of payments may be symptomatic of a problem, and to consider what this analysis implies currently for the United States. With regard to the last, we will suggest that concern about the U.S. trade deficit has been overstated. Indeed, a trade deficit can be indicative of a healthy and strongly growing economy.

THE BALANCE OF PAYMENTS ACCOUNTS

The balance of payments accounts are a record of transactions between domestic residents and the rest of the world over a specific period of time. Like any double entry bookkeeping system, the balance of payments accounts must balance. There is nothing mysterious about this, nor does it involve any statement about how the world works.

The simplest form in which the balance of payments accounts can be expressed is as follows:

\[ CA + K + F = 0, \]

where CA is the current account balance, K is net non-official capital flows and F is official reserve financing. These items are defined in such a way that they must sum to zero. Let us consider each of them in turn.

Current Account

The current account has two major components. These are the trade balance and the services or "invisibles" balance. The former, which generally gets the most attention, is the difference between the value of goods exported and the value of goods imported. These exports and imports are of physical objects which, in principle, could be observed crossing the border. In contrast, "invisibles" are services for which international payments are made but that do not

\[ \text{Because of measurement errors, the actual accounts add in a "statistical discrepancy" which when included in (1) ensures balance. The reason we say that they must balance, however, is not a statement about the accuracy of the statistics. The current and capital account (including official balance) are defined to be equal and opposite. Think of the current account as the excess of income over spending. The capital account is then merely net saving, which is equal to income minus spending. If you measure saving as negative and the excess of income over spending as positive, they will obviously add up to zero.} \]
Invisible import for the United States and an invisible export for that insurance contract would represent an invisible export. For example, if a New York shipping company were to insure a cargo with Lloyds of London, the purchase of such insurance would involve the direct transfer of a physical product. For example, if a New York shipping company were to insure a cargo with Lloyds of London, the purchase of insurance would represent an invisible export.

Invisibles take many different forms. Two examples are worth mentioning in addition to such financial services as insurance and banking. First, if a nation has either assets or liabilities overseas, the net payment of interest or dividends is measured as an invisible import or export. A positive net return on foreign assets is counted as an invisible export, because it generates an inflow of payments into the economy just as an export of goods does. Second, international tourism is counted as part of the invisible component of the current account. If U.S. citizens spend more on overseas trips than foreigners spend on U.S. vacations, it is measured as an invisible net import in the U.S. balance of payments.

**Non-Official Capital Account**

The capital account of the balance of payments measures the change in net indebtedness between the domestic economy and the rest of the world. It is important to get this clear, as there is sometimes confusion about what the capital account contains. It does not involve imports and exports of capital goods, such as machine tools and computers. These are all physical goods, and their import and export are therefore counted in the trade account. The capital account involves the transfer of financial claims of various kinds. These claims are referred to as “capital” because they represent claims to interest or dividend payments and, in the case of company shares, do involve ownership of underlying real assets.

The terminology commonly used to describe the capital account is rather confusing when it is related to the way in which capital account items are measured. In the current account of the balance of payments, goods leaving the country is measured as a plus item. In the capital account, however, what is generally called a capital “outflow” is measured as negative. Only the terminology here is confusing, however; accounts are quite logical. What we mean by a capital outflow is that domestic residents are buying foreign assets. In other words, they are “importing” foreign shares, titles or securities. Thus, all purchases of foreign goods, securities (stocks, bonds, bills) or any other asset are measured as negative (imports), and all sales to foreigners are measured as positive (exports) irrespective of whether they are goods sales or asset sales.

In principle, the capital account of the balance of payments measures the change in the net asset/liability position between the home economy and the rest of the world. We say “in principle” because there is one respect in which this is not correct. The capital account measures the value of the net flow of financial instruments (stocks, bonds, bills, etc.) that passes between domestic and overseas residents. But the external indebtedness of an economy changes not just as assets change hands. It also changes as a result of changes in values of assets that have not changed hands. For example, U.S. residents may own shares in Rolls Royce which rise in value. This capital gain (or loss) element of the external asset/liability position is not measured as part of the balance of payments accounts until it is realized by an asset sale. Only the flow of financial claims is included.¹

**Official Balance**

The final item in the balance of payments accounts is the balance for official financing. This comprises changes in the official foreign exchange reserves of the domestic economy. These reserves are mainly claims against foreign governments (or central banks), for example, Fed holdings of Deutsche mark. For most countries, reserves are held as a means of intervening in foreign exchange markets to support the value of the domestic currency.² This item is a special official sector component of the capital account. It is treated separately for historical reasons associated with the fixed exchange rate system which operated almost worldwide from World War II until 1973.³ Under a freely floating exchange rate regime, the official financing balance is always zero. If F in equation 1 is zero, clearly, CA and K must be equal and of opposite sign.

**U.S. Balance of Payments**

Table I shows the U.S. balance of payments for 1986. It shows a current account deficit of a little over $141 billion. The current account is made up of items 1 and 2. The capital account surplus of $117 billion is shown in lines 3 and 4. Changes in U.S. official reserves are shown in line 5. There was a very small fall of $0.312 billion in 1986 (a plus sign indicates a decline in holdings of foreign assets). This indicates that the U.S. authorities intervened little during 1986 as a whole.

---

¹Some have claimed that the United States has become a net debtor vis-a-vis the rest of the world. This claim ignores the capital gains on U.S.-owned foreign assets. In reality, the United States is likely still to have positive net external assets.

²See Balbach (1978).

³See Batten and Off (1983).
Table 1
U.S. Balance of Payments: 1986
(millions of dollars)

<table>
<thead>
<tr>
<th>Item</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Merchandise trade</td>
<td>$144,399</td>
</tr>
<tr>
<td>2) Invisibles net</td>
<td>+2,987</td>
</tr>
<tr>
<td>Balance on current account</td>
<td></td>
</tr>
<tr>
<td>(1 plus 2)</td>
<td>$141,352</td>
</tr>
<tr>
<td>3) Change in U.S. assets abroad</td>
<td>-96,294</td>
</tr>
<tr>
<td>(increase -)</td>
<td></td>
</tr>
<tr>
<td>Balance on capital account</td>
<td>+213,387</td>
</tr>
<tr>
<td>(3 plus 4)</td>
<td></td>
</tr>
<tr>
<td>4) Change in foreign assets in U.S. (increase +)</td>
<td>+117,093</td>
</tr>
<tr>
<td>(3 plus 4)</td>
<td></td>
</tr>
<tr>
<td>Statistical discrepancy</td>
<td>-312</td>
</tr>
<tr>
<td>(1 + 2 + 3 + 4 - 5)</td>
<td></td>
</tr>
<tr>
<td>$23,947</td>
<td></td>
</tr>
</tbody>
</table>

NOTE: The merchandise trade balance is exports minus imports.
The invisibles balance is the sum of: net military transactions; net investment income; other service transactions; net remittances; pensions and other transfers; and U.S. government grants (non-military).

SOURCE: U.S. Department of Commerce.

Thus, the dominant picture is one of U.S. residents buying more goods and services overseas than foreign residents are buying from the United States and of foreigners increasing their net holding of claims against the United States.

Notice, however, that there is a fairly large statistical discrepancy. The presence of this discrepancy indicates that the data do not include some trade and/or capital flows. While it is impossible to say where the inaccuracies arise, it is often presumed that the greatest errors are likely to be in the capital account, primarily because asset transfers are more difficult to keep records on. If the data had no omissions, then the current and capital accounts (including official flows) would add to zero.

It is not obvious at first glance why the current and capital accounts must offset each other exactly. What would happen if they did not? Suppose for example, that at current exchange rates a country is running a current account deficit but its planned net capital flows are zero. This means that the country is trying to spend more on imports than foreigners are willing to spend on its exports. This will produce an imbalance in its foreign exchange market. Attempted sales of domestic currency (for foreign currency) will exceed attempted purchases. The market value of the currency will fall until the quantity of the currency demanded is equal to that supplied. At this point, either the current account has adjusted so that it is no longer in deficit, or the net export of assets induced as assets in the country became cheaper, through domestic currency devaluation, and thus more attractive to foreigners, and prices of foreign assets became higher and hence less attractive to U.S. citizens is just equal to the current account deficit. Thus, the exchange rate will adjust to ensure that the current and capital accounts are exactly offsetting.

There is nothing magical about this outcome. The end result is the same for any individual. If you spend more than your income, you must borrow or sell the equivalent value of your assets to cover the difference; if you spend less than your income, you must inevitably acquire increased claims on someone else. Similarly, a nation that runs a current account deficit must either borrow from abroad or sell off some of its assets, whether these assets are domestic or foreign. Likewise, a current account surplus must be associated with either an increase of claims on foreigners or a reduction of previous borrowings.

Another implication of the definition of balance of payments is the following identity:

\[ 12) \quad \text{CA} = \text{GNP} - \text{GDE}. \]

The current account surplus (or minus the current account deficit) is equal to gross national product minus gross domestic expenditure. This identity shows that the current account of the balance of payments is the difference between the value of what the nation produces and what it spends. The former (GNP) can also be thought of as the value of the nation’s gross income. Identity (12) is useful because it makes clear that any nation that spends more than it produces will have a trade deficit. The interesting question, of course, is whether such an imbalance is good or bad.

**WHAT MAKES THE CURRENT ACCOUNT BALANCE A PROBLEM?**

The nature of what is usually termed a balance of payments problem varies considerably, depending

\[ \text{See Chrydsta (1984).} \]
upon whether the country in question has a fixed or a floating exchange rate regime. The problem produced by a deficit on the current account can be most acute if the nation is maintaining a fixed exchange rate regime. In this case, "the problem" is felt directly by the central bank.

Maintaining a fixed exchange rate vis-a-vis one or more countries requires the pegging nation's central bank to hold foreign exchange reserves with which to intervene in the foreign exchange market. This intervention can be necessary to stop the exchange rate from moving in either direction. Suppose, for example, that the country has a current account deficit and no desired net private capital flows. In order to maintain the existing exchange rate, the central bank must sell foreign exchange for its domestic currency. Whether the origin or source of the net supply of domestic currency in foreign exchange markets is from the current or capital account side of the balance of payments is irrelevant. The domestic currency value of reserves sold in a particular period is the official financing balance, F, in equation 1. Because it involves the sale to foreigners of a domestically held asset, a net loss of reserves is measured as positive in the balance of payments accounts.

Under a fixed exchange rate regime, exchange rate pressure poses a problem if the central bank in question starts to run out of foreign exchange reserves. This possibility makes the problem worse because holders of the domestic currency, fearing a devaluation, will try to buy foreign currency. Speculative sales of the domestic currency in foreign exchange markets force the central bank to sell even more foreign exchange reserves. Inevitably, the nation must either devalue its currency or introduce measures to cut domestic spending (including spending on foreign goods). This action is unavoidable; otherwise, the central bank will run out of foreign exchange reserves.

This describes the nature of most balance of payments crises experienced by countries attempting to maintain fixed exchange rates in the 1950s and 1960s. It is worth noting, however, that the United States under the postwar "Breton Woods" regime was not the same as other countries. All other countries in the system pegged their currencies to the dollar and held dollar reserves for this purpose. The United States, therefore, did not need to support its own exchange rate and, in fact, did not hold significant reserves of foreign currency during this period.

Since the spring of 1973, when all the major industrial countries moved to a floating exchange rate regime (the United Kingdom had floated in June 1972), the nature of balance of payments problems has changed. Under a floating exchange rate system, a central bank does not have to use its foreign exchange reserves to finance a deficit in the non-official part of the balance of payments: in fact, there will be none. In equation 1 above, the term F becomes zero. Instead of central bank intervention, the exchange rate moves to assure that the current account and the capital account sum to zero on their own.

WHY WORRY ABOUT THE TRADE BALANCE?

Concern about the state of the trade balance has a long history. It is useful to put this concern in historical context, as it leads naturally to the analysis of when such concern is justified.

In the following discussion, we take it as given that trade itself is beneficial, a point not clearly established until Ricardo's famous demonstration published in 1817. There was, however, some connection historically between the case against trade deficits and the understanding of why trade in general was a good thing. Only when the gains from trade were properly understood could people begin to make sensible assessments of the cause and effect of trade deficits.

The context in which the early debates took place was an international economy in which payments for external trade were largely made in precious metals, especially gold. The effect of running a trade surplus was that a nation would accumulate gold. In many

---

7The exception to this is when a currency is depreciating at a fast rate. This is a symptom of acute internal problems normally associated with hyperinflation.

8The system was named after the place in New Hampshire where the final negotiations setting it up were held in July 1944.

9The U.S. authorities agreed to convert dollars into gold at $35 per ounce. This commitment was abandoned for all but official holders in March 1968 and for official holders in August 1971. See Batten and Ott (1983) for evidence on exchange market intervention.

10Note that even today the majority of small countries peg their exchange rates to either a major currency or a weighted basket of currencies. Reserve shortages still may cause acute problems for them.

11In fact, none of the major currencies are floating freely. All the major central banks have intervened from time to time to influence exchange rates. Intervention to support the dollar has been especially heavy since the "Plaza Accord" of September 1985.
people's minds, the accumulation of gold itself became the object of trade: trade surpluses were "good" and trade deficits "bad." Trading in order to build up gold holdings became known as mercantilism.

Mercantilism was criticized by several eminent writers, including David Hume (1752), who showed that a continuing trade surplus was unattainable. An existing trade surplus, he noted, produces an inflow of gold. Because gold is a form of money, the quantity of money in the country rises. This, in turn, produces a rise in prices, which continues as long as more gold flows in. As the country's goods become more expensive relative to those produced overseas, however, fewer will be bought, eventually eliminating the trade surplus. 12

Some years later, David Ricardo (1817) used this demonstration to show why trade deficits occurred. His answer to this question brings us directly to our central point: trade deficits can result from a variety of sources, not all of which are "bad."

**RICARDO, THORNTON AND TRADE DEFICITS** 13

Ricardo argued that a trade deficit was the inevitable consequence of prices in the deficit country being "too high." These prices, in turn, were produced by excessive prior monetary expansion from domestic sources that were unrelated to prior trade surpluses. 14 He argued, in other words, that excessive monetary expansion was not only a sufficient condition for a trade deficit to occur, it was also a necessary condition and vice versa for trade surpluses.

This describes what happened in many countries during the Bretton Woods regime. While this sequence of events portrays a common cause of trade deficits, however, it is not the only cause. In the 65-year period between 1830 and 1895, the United States had a current account deficit in almost every year; there were only 13 years in which a surplus was recorded. Yet this was not a period of sustained inflation. Indeed, it was a period of rapid and prolonged economic growth. There is thus at least one counterexample — and a major one — to Ricardo's generalization. How can this be explained?

At the time Ricardo was writing, his claim was disputed, most notably, by Henry Thornton. Thornton argued that, although prior excessive money expansion was indeed sufficient to produce a trade deficit, it was not a necessary condition for a trade deficit.

Thornton distinguished between trade deficits arising from real causes and those arising from excessive money creation. The former can occur because individuals in a country want to spend more than their current income, that is, they wish to reduce their net financial wealth or increase their net indebtedness. 15 In terms of equation 2 above, anything that causes domestic spending to exceed output will produce a trade deficit.

Of course, the balance of payments deficit from this cause can not persist forever. It will disappear when individuals have reached their new lower desired wealth level; in the same manner, a trade deficit produced by excess money creation will end when the excess money has been dispersed overseas (or deflated by higher prices). 16

In summary, a trade deficit can be produced not just by excess monetary expansion, but by dissaving. 17 Both of these will produce deficits that are temporary; however, these deficits will be eliminated eventually by different mechanisms. Dissaving and the associated decline in financial wealth can be produced by several factors; examining some major ones helps to understand the current U.S. situation.

---

12Hume, although dealing explicitly with the mercantilist argument, dealt implicitly with the notion that an export surplus is necessary for growth. Since a perpetual export surplus is impossible, if an export surplus were essential for growth, growth would have stopped. It did not, however, and to date has not.

13An extensive discussion of the ground covered in this section can be found in Perelman (1986).

14The issue of domestic bank notes partially backed by gold was a topic of controversy between the "currency" and "banking" schools through the 19th century in Britain.
WHY SHOULD THERE BE DISSAVING?

In order to discuss the possible sources of dissaving in the domestic economy, it is convenient to set out another identity:9

\[ (3) \quad CA = (S - I) + (T - G). \]

This shows that the current account surplus must be equal to the excess of private saving over private investment \((S - I)\), plus the government budget surplus \((T - G)\). In other words, the surplus for the economy as a whole can be broken down into the private sector surplus plus the public sector surplus. This classification suggests possible directions in which to look for causes of the trade deficit: a fall in private saving, a rise in private investment or an increase in the government budget deficit.

A fall in private saving must be associated with an increase in consumption relative to income. This could happen if there were a temporary fall in income due, for example, to a crop failure or a natural disaster. It is well established that, at times when income is abnormally low, people attempt to maintain their consumption patterns by dissaving. If the nation as a whole does this, it will necessarily involve a trade deficit. It should be emphasized that, while crop failures or other natural disasters are unfortunate, the ability to adjust to these events by dissaving and thus importing goods from abroad is preferable to reducing domestic consumption. In extreme cases, the choice may be between running a trade deficit and starvation. While natural disasters can explain some trade deficits, it is unlikely to explain the U.S. deficits in the 1980s. After all, this has been a period of fairly steady income growth.

The second alternative suggested by identity (3) is a rise in private investment, caused by an expected rise in the productivity of domestic capital (relative to that overseas). This alternative is an extremely healthy sign for the domestic economy. It indicates that the expected profitability of investment was such that firms were prepared to borrow in order to finance the higher investment. If private investment exceeds private saving (for a balanced government budget), the private sector must borrow from overseas. We have seen already that net borrowing from overseas implies a current account deficit in the balance of payments.

If overseas-financed growth in private investment lies behind the trade deficit, we have to be careful in interpreting the statement that the trade deficit is associated with dissaving or a reduction in wealth. It is true that the private sector will be increasing its net financial liabilities (or reducing net financial assets). At the same time, however, it is converting those liabilities into real capital. The return on that real capital is expected to be greater than the cost of the borrowing. Hence, this provides the basis for income and wealth growth in the future and, presumably, explains why the United States had sustained trade deficits throughout the second half of the 19th century. Rapidly growing countries that attract capital from overseas typically will have trade deficits.

The final possibility is that the current account deficit reflects the government budget deficit. Obviously, if private saving and investment were equal, the budget deficit and the current account deficit would be equal. We shall not pursue the question of whether the budget deficit is "good" or "bad" for the economy. Assuming that the budget deficit represents the deliberate choice of policymakers, however, it follows that the associated trade deficit must be preferred to the alternatives.

Thus, we have seen that a rising current account deficit must be associated with either a rise in investment relative to saving (or fall in saving relative to investment) or a rise in the budget deficit of the government. We already had seen that current account deficits could result from excessive monetary expansion, a case that is consistent with identity (3): the attempt to spend the excess money will result in either a fall in \(S-I\) (higher consumption, lower saving or higher investment) or a fall in \(T-G\) (more government spending relative to taxes).

The Evidence for the United States

We now look at the possible causes of the U.S. current account deficit. First, we consider the argument, favored by Ricardo, of fast monetary growth associated with high domestic inflation. At first sight, this appears a likely possibility. Monetary growth accelerated after 1982 (chart 1) at the same time as the current account plunged into deficit (chart 2). However, U.S. inflation fell (chart 3) and remained consistently below the OECD average during this period. Also, both the real and effective exchange rates appreciated strongly until 1985. The inflation and exchange rate behavior are signs of monetary tightness, not mone-

---

9This can be derived as follows: \(\text{GNP} = C + I + G + CA\) from the expenditure accounts. It is also true that \(\text{GNP} = C + S + T\) from the income accounts. So \(I + G + CA = S + T\) and \(CA = (S - I) + (T - G)\).
Only the high U.S. money growth in 1986 looks consistent with Ricardo's explanation: both the real and nominal exchange rates fell during 1986. The increase in the trade deficit in 1986, however, was small. Hence, little weight can be attached to the monetary explanation of the trade deficit. Indeed, why the rapid money growth of 1982–86 did not create inflation is still something of a mystery. There was, over this period, a significant decline in the velocity of circulation, which means that the extra money balances were willingly held rather than spent domestically.21

A much more plausible story emerges from a plot of the private and public sector surpluses (chart 2). Notice that we show here I – S rather than S – I, because it is easier to see its correspondence with T – G. Before 1982, the relationship between the public sector deficit and the private sector surplus was remarkably close. As a result, current account deficits and surpluses generally were small. After 1982, however, the

20It is possible that the Mundell model referred to above is relevant here. This predicts that monetary tightness causes capital inflows, a currency appreciation and a current account deficit. We think this unlikely to be relevant here. There is no clear evidence of sufficient monetary tightening over the entire 1981–86 period to explain what happened. More importantly the same outcome is predicted from the Mundell analysis as resulting from fiscal expansion. Hence monetary neutrality combined with fiscal expansion would be sufficient. It is the latter which seems to us to dominate in this case.

CONCLUSION

A trade deficit arises when a country buys more from overseas than foreigners buy from it. The counterpart of a trade deficit in the balance of payments accounts is an increase in borrowings (or reduction in net lending) from the rest of the world. Trade deficits could result from inflationary domestic monetary policies; there is no evidence, however, that such policies are the cause of the current U.S. trade deficit. In general, a trade deficit must be associated with some combination of private and public sector deficits. Until 1982, the budget deficit was approximately financed by private sector surpluses. The present situation, however, is the inevitable result of the combination of a budget deficit and high investment relative to private saving.

REFERENCES


Hume, David. Political Discourses (Edinburgh, 1752).
Chart 3
U.S. and OECD Inflation


