

The Mystery of the Excess Trade (Balances)

By DONALD R. DAVIS AND DAVID E. WEINSTEIN*

Bilateral trade deficits are a perennial policy issue. Former Deputy Assistant U.S. Trade Representative for Japan and China, Merit Janow (1994 p. 55), notes that during the first George Bush administration, “High deficits coupled with the continuing allegations from U.S. business interests about the closed nature of the Japanese market were resulting in serious domestic political pressures for improved access to the Japanese market.” Recently Robert C. Feenstra et al. (1998 p. 1) made similar comments vis-à-vis China: “Some analysts have interpreted the large U.S.–China bilateral trade deficit as *prima facie* evidence of unacceptably high levels of protectionism in China, and have advocated stringent entry conditions for China’s admission into WTO.”

Given the policy salience of bilateral trade deficits, it is peculiar that no one has ever examined them empirically for a broad set of countries. One reason for the scant study is that economists are naturally (and sensibly) loath to accept the terms of the policy debate, which considers bilateral trade deficits *ipso facto* harmful. A second reason is that economists believe there may be very natural explanations for bilateral imbalances. One such explanation finds its origins in macroeconomic identities that equate current-account deficits to an excess of investment over saving. From this, it may be argued that bilateral imbalances will arise naturally in trade between countries in aggregate surplus and those in aggregate deficit. Indeed, this is the principal explanation that the profession has given policymakers, and it forms the foundation of many U.S. bilateral trade initiatives such as the Structural Impediments Initiative and the Framework talks. Janow (1994 p. 55) observes that “there was (and is) little disagreement among economists that the causes of

large aggregate *and bilateral* deficits are largely attributable to macroeconomic factors” [italics added]. A second account may rely on what may be termed “triangular trade,” in which cross-country differences in the patterns of demand and supply mean that a country will run bilateral deficits with those countries that are unusually important suppliers of the goods for which the deficit country happens to be an unusually strong demander.

In this paper, we use the canonical “gravity model” of bilateral trade to form predictions about bilateral trade balances. We develop two key variants of the model, in which bilateral trade imbalances arise due to aggregate macroeconomic imbalances or due to “triangular trade,” and we implement these empirically for a broad set of countries. Our results paint a dismal picture. The central explanations that economists provide to explain bilateral balances perform miserably. There are two key failures. First, actual bilateral trade imbalances are much larger than those predicted; there is a “mystery of excess trade balances.” Second, even after we allow for both macroeconomic imbalances and idiosyncrasies in the structure and levels of demand and production, the models perform poorly in explaining bilateral trade balances. These failures of economists’ standard explanations of bilateral trade imbalances require that we move beyond the simple gravity framework to consider alternative explanations: homogeneous goods, highly specialized intermediates, and the role of policy.

I. Theory

The dominant intellectual paradigm for understanding bilateral trade patterns is the so-called gravity model of trade. This then also seems an appropriate starting point for making sense of bilateral trade imbalances. We will start with a very simple model that ignores trade frictions, incorporating these explicitly only when we turn to empirics.

Let X_c be GDP in country c , and let s_c be its share of world spending. Let world GDP be

* Department of Economics, Columbia University, 420 West 118th Street, New York, NY 10027, and National Bureau of Economic Research. Address correspondence to Davis. We especially thank Josep Vilarrubia for research assistance and Robert Feenstra for comments.

$X_w = \sum_c X_c$. Assume that consumers in each country have identical homothetic preferences, that trade is perfectly free, and that consumers perceive goods in different countries to be distinct goods. Then we can write exports from c to c' as:

$$(1) \quad E_{cc'} = s_{c'} X_c.$$

This also allows a very simple statement of the bilateral trade balance:

$$(2) \quad T_{cc'} \equiv E_{cc'} - E_{c'c} = s_{c'} X_c - s_c X_{c'}$$

If we let TD_c be country c 's aggregate trade deficit, then country c 's share of spending can be written as:

$$(3) \quad s_c = \frac{X_c + TD_c}{X_w}.$$

If we let td_c be country c 's aggregate trade balance scaled by world GDP, then the bilateral trade balance becomes:

$$(4) \quad T_{cc'} = s_{c'} X_c - s_c X_{c'} = \frac{X_{c'} + TD_{c'}}{X_w} X_c - \frac{X_c + TD_c}{X_w} X_{c'} = td_{c'} X_c - td_c X_{c'}.$$

That is, *bilateral* trade imbalances arise exclusively as a result of *aggregate* trade imbalances. A special case of this would be when all countries run balanced trade in the aggregate (i.e., $td_c \equiv 0 \forall c$). In this last case, all bilateral balances would likewise be zero. A perspective such as that embodied in equation (4) could rationalize a claim that it is natural for a country running an aggregate trade deficit, such as the United States, to run large bilateral deficits with a country running aggregate trade surpluses, such as Japan.

So far our theory has ruled out the possibility that the structure of demand may vary across countries, which becomes relevant when the industrial structure of countries likewise differs. This is the setting in which "triangular trade" could be important. To make sense of this, let α_{ic} be country c 's share of its own spending devoted to industry i and let ϕ_{ic} be the share of

its GDP arising in sector i . Then it is straightforward to show that the bilateral trade balance between c and c' is given by:

$$(5) \quad T_{cc'} = \sum_i \left[\alpha_{ic'} \left(\frac{X_{c'} + TD_{c'}}{X_w} \right) \phi_{ic} X_c - \alpha_{ic} \left(\frac{X_c + TD_c}{X_w} \right) \phi_{ic'} X_{c'} \right]$$

In such a case, bilateral trade imbalances may arise, as before, due to the countries' aggregate imbalances. The new forces come from the interaction of differential structure in demand (α_{ic}) and supply (ϕ_{ic}). The new forces are easier to see if we restrict the aggregate balances to be zero, which allows the following simplification:

$$(6) \quad T_{cc'} = \sum_i [\alpha_{ic'} \phi_{ic} - \alpha_{ic} \phi_{ic'}] \left(\frac{X_c X_{c'}}{X_w} \right).$$

Because α_{ic} and ϕ_{ic} are shares, equivalence across c and c' either in demand or production structure would have the consequence that aggregate balance ensures bilateral balance. However, when countries differ in their demand structures and also differ in their production structures, bilateral imbalances are quite natural, even if the countries are in aggregate balance. Obviously to be in aggregate balance requires that bilateral deficits with some countries be offset with surpluses with other countries, hence rationalizing the triangular-trade explanation for bilateral trade deficits.

To summarize, we have developed a theory of bilateral trade balances within the paradigmatic model of bilateral trade, the gravity model. If all countries run balanced trade in the aggregate and either the structure of demand or of production is common across countries, then all bilateral balances are predicted to be zero. Aggregate trade imbalances alone suffice to give rise to bilateral trade imbalances, as per equation (4). If there are aggregate trade imbalances and differences in the structure of demand or production, then the bilateral balances are given by equation (5). If aggregate trade is balanced, bilateral imbalances can still arise if both demand *and* production patterns vary across countries, as in equation (6).

II. Empirics

The empirical question we examine can be stated simply. How successful is the gravity model and simple amendments, as embodied above in equations (1)–(6), in explaining actual bilateral trade balances? We begin with the simpler model, based on equation (4), which traces bilateral imbalances to macroeconomic imbalances and then move on to consider triangular trade, as in equation (5).

Our data include exports and output for a sample of 61 countries and 30 industries at the three-digit ISIC level for manufacturing, agriculture, and mining. Sources for the data are Feenstra et al. (1997), United Nations (1997), United Nations Industrial Development Organization (1999), and Shang-Jin Wei (1996), and a more detailed description is available from the authors upon request. The key variables are standard in the gravity literature. The dummy variable FTAEC is unity if both members of a country pair were part of NAFTA or the EC. REMOTE is an inverse distance-weighted average of rest-of-world GDP's. $DIST_{cc'}$ is the bilateral distance between countries c and c' , and $ADJ_{cc'}$ is an indicator variable for a common border. We can write a gravity specification that controls for these additional factors as follows:

$$(7) \quad \ln E_{cc'} = \beta_0 + \beta_1 \ln(s_c X_c) \\ + \beta_2 \ln(DIST_{cc'}) \\ + \beta_3 \ln(REMOTE_c) \\ + \beta_4 \ln(ADJ_{cc'}) \\ + \beta_5 \ln(FTAEC_{cc'}) + \varepsilon_{cc'}$$

We begin by estimating equation (7) using aggregate bilateral exports as our dependent variable, GDP as our proxy for X , and GDP plus the current account as our proxy for s_c . The estimation is based on a Tobit procedure. The fits and coefficient estimates are entirely conventional. We then take the exponential of the fitted values to calculate estimated bilateral balances, $\hat{E}_{cc} - \hat{E}_{c'c}$. We plot these against the actual imbalances $E_{cc} - E_{c'c}$ in Figure 1.

These results may be interpreted as a simple test of the macroeconomic balance approach to

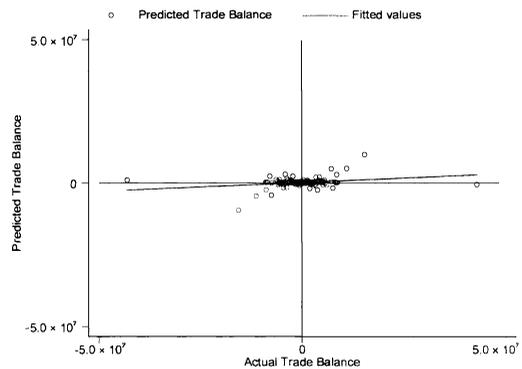


FIGURE 1. ESTIMATED VERSUS ACTUAL TRADE IMBALANCES

bilateral trade balances. The results reveal an interesting feature of the data. Had the model simply not fit well, one would have expected to see the predicted bilateral balances exhibit a similar variance to that of the actual balances. Instead we see that, with a few exceptions, our model predicts balances that are an order of magnitude smaller than actual imbalances. The ratio of the variance of predicted balances to actual balances is just 0.05. The macroeconomic approach to bilateral trade balances predicts the correct sign of the bilateral balance only 54 percent of the time—barely better than a coin flip. Regression evidence confirms the visual impression: The coefficient of fitted imbalances on actual trade imbalances is 0.06 and the R^2 value is 0.07. If we control for outliers by running a median regression, the performance of the model deteriorates further. Variation in macroeconomic balances just do not explain bilateral trade balances.

One hint at the problems in the macroeconomic balance approach comes from examining the source of U.S. imports. If macroeconomic balances were the entire story, then, controlling for distance, every country should send the same share of their exports to the United States. However, this is not at all what the data indicate. Consider the patterns of exports from several East Asian countries. China sent 9 percent of its exports to the United States, while Hong Kong and Japan sent 23 percent and 37 percent, respectively. Similar stories can be told for many bilateral trade patterns. This underscores the notion that actual bilateral export flows are far more variable than what one might expect by

looking at aggregate absorption and production. This leads to us consider our second standard approach to understanding bilateral trade.

An obvious explanation for the variability in export shares is complementarities between industrial production and demand patterns. If Japan produces goods that the United States wants, but China does not, then this can naturally give rise to variations in the deficit. In order to test this hypothesis, we added industry subscripts to our gravity equation and estimated it separately for each of 28 three-digit ISIC manufacturing categories, as well as for agriculture and mining. Our new specification is

$$(8) \quad \ln E_{icc'} = \beta_{i0} + \beta_{i1} \ln(s_{ic'} X_{ic'}) \\ + \beta_{i2} \ln(\text{DIST}_{cc'}) \\ + \beta_{i3} \ln(\text{REMOTE}_c) \\ + \beta_{i4} \ln(\text{ADJ}_{cc'}) \\ + \beta_{i5} \ln(\text{FTAEC}_{cc'}) + \varepsilon_{icc'}$$

Here, X_{ic} is defined as industry output, and the share of absorption is defined as output less net trade in that sector. Once we estimated this equation, we then took the exponential of the fitted values to calculate estimated sectoral bilateral balances, $\hat{E}_{icc'} - \hat{E}_{ic'c}$. We then summed these across all sectors and plotted them against the actual imbalances $E_{cc'} - E_{c'c}$. This specification employs much more information in order to predict bilateral flows. Not only do we allow absorption and output to vary across countries, but we also now use 180 parameters instead of five.

We could plot these results, and they would look slightly better than those in Figure 1. This is not really surprising given the great increase in parameters and explanatory variables. What is perhaps more surprising is that this does not eliminate the excess-trade-balance phenomenon. The variance of actual trade balances is still 4.5 times larger than the variance of predicted balances. Moreover, regressing predicted bilateral balances against actual balances reveals that the model's prediction picks up very little of the variation in the actual balances.

Once again, when we look within sectors, the same problem that we saw in the aggregate-

flows data emerges: there is much more variability in export flows than one might expect even given the cross-country differences in demand and production patterns. Consider the case of the third-largest export sector, electrical machinery. The United States absorbs 19 percent of the world supply of this industry. However, only 5 percent of Chinese electrical machinery was shipped to the United States while the corresponding numbers for Hong Kong, Japan, and the Philippines stood at 25, 34, and a whopping 54 percent.

III. Conclusion

Bilateral trade balances are an important source of frictions in international trade relations, so it is important to understand their provenance. In this paper, we provide an empirical examination of two key theories—one based on macroeconomic balances and the other based on triangular trade. The theories perform poorly in explaining bilateral trade balances. Actual bilateral trade balances are vastly larger than those predicted by theory, a result that may be termed the "mystery of the excess trade balances." The poor performance of the model is likewise confirmed by the poor overall fits and the very weak ability even to predict the sign of the bilateral trade balances. The failure of these models to explain actual bilateral trade balances does not imply that bilateral protection is the source of these imbalances. However, it should force international economists to reflect on the deficiencies of the gravity framework in this regard and to consider alternative explanations (possibly including bilateral protection) for understanding these mysteries.

REFERENCES

- Feenstra, Robert C.; Lipsey, Robert E. and Bowen, Harry P. "World Trade Flows, 1970–1992, with Production and Tariff Data." National Bureau of Economic Research (Cambridge, MA) Working Paper No. 5910, 1997.
- Feenstra, Robert C.; Hai, Wen; Woo, Wing T. and Yao, Shunli. "The US–China Bilateral Trade Balance: Its Size and Determinants." National Bureau of Economic Research (Cambridge, MA) Working Paper No. 6598, 1998.

Janow, Merit E. "Trading with an Ally: Progress and Discontent in U.S.-Japan Trade Relations," in Gerald L. Curtis, ed., *The United States, Japan, and Asia: Challenges for U.S. policy*. New York: Norton, 1994, pp. 53-95.

United Nations. *United Nations statistical yearbook*, 41st issue. New York: Statistics Division, Department for Economic and Social Information and Policy Analysis, United Nations, 1997 [CD-ROM].

United Nations Industrial Development Organization. *UNIDO industrial statistics Database 1999*, 3-Digit Level of ISIC Code. New York: Statistics and Information Networks Branch, United Nations Industrial Development Organization, 1999 [on diskette].

Wei, Shang-Jin. "Intra-national versus International Trade: How Stubborn are Nations in Global Integration?" National Bureau of Economic Research (Cambridge, MA) Working Paper No. 5531, 1996.