To our readers:

We begin the year 2000 with a new design for the *Review*. We adopted a new design five years ago to improve readability and to enhance reproduction, particularly for classroom use. We think this new design takes us a step further in that direction.

The *Review* continues to serve a diverse audience—as we have for decades—from the interested layman to graduate students in economics and policymakers. Our goal is to produce useful and relevant information for all of our readers. We will continue to publish a wide array of articles, ranging from essays on policy issues for general readers to technical treatments of economic issues. As in the past, our editorial policy is to place articles written for the general reader first in each issue.

We continue to appreciate hearing from readers with questions and comments on our articles. We also welcome comments on our new design.

William T. Gavin
Editor
January 3, 2000

Cletus C. Coughlin is a vice president and associate director of research and Patricia S. Pollard is an economist and research officer at the Federal Reserve Bank of St. Louis. Heidi L. Beyer provided research assistance.

State Exports and the Asian Crisis

Cletus C. Coughlin and Patricia S. Pollard

INTRODUCTION

Real merchandise exports from the United States to East Asia fell by 12 percent during 1998 as the Asian crisis reduced demand in the region.[1] These markets accounted for about 30 percent of U.S. exports prior to the crisis. Given this market share, the 12-percent drop in merchandise exports to East Asia would have resulted in a 4-percent drop in total U.S. merchandise exports, absent any changes in export sales elsewhere. Because merchandise exports account for 10 percent of U.S. output, the 4-percent decline in exports would have resulted in a 0.4-percent decline in U.S. output. This decline did not occur because exports to the rest of the world increased and, more importantly, strong U.S. domestic demand offset the negative effects of the trade shock.

Despite the overall benign effect on the U.S. economy, the Asian crisis produced numerous microeconomic effects. In a recent article in this *Review*, Pollard and Coughlin (1999) estimated the effects of the decline in exports to East Asia on an industry-by-industry basis.[2] Exports to East Asia fell during 1998 for 25 of the 26 industries studied. These declines ranged from 35 percent for the metallic ores and concentrates industry to 3 percent for the food and kindred products industry. The one industry whose exports to East Asia rose was the transportation industry, driven primarily by increases in aircraft exports to China and Taiwan.

In the absence of offsetting effects, the relevance of these declines in exports for the sales of each industry depended on the importance of East Asian markets for each industry's production. Taking account of this, Pollard and Coughlin (1999) calculated that the nonelectrical machinery industry was the one affected most adversely by the decline in exports to East Asia. The growth rate of gross output in that industry would have been 1.8 percentage points higher if exports to East Asia had not declined. Of those industries whose exports to East Asia fell during 1998, the printing and publishing industry was affected the least, primarily because East Asia was not an important outlet for its production.

Because industry composition varies across states, the different export effects across industries suggest that individual states may have been affected to varying degrees by the Asian crisis. Furthermore, the geographic pattern of trade varies across states indicating that those with a high proportion of exports going to East Asia were more likely to have been affected by the crisis than others. Although the effects of the Asian crisis on specific states have been addressed in a number of studies, to date, no comprehensive study of the export effects across states has been published.[3] One study providing a state-by-state analysis was released during early 1998 by the Department of Commerce and the Treasury Department (1998). The goal of that study, however, was to predict the states that were most likely to be affected, rather than to calculate the actual effects.

This article provides an additional perspective on the microeconomic effects of the Asian crisis focusing on manufacturing sectors of individual states. We begin with an overview of state-level manufacturing exports. Next, we examine the changes across states in manufacturing exports to East Asia. We then examine the importance of the East Asian markets to each state. Combining estimates of the change in a state's exports to East Asia with the assessment of the importance of these exports to the manufacturing sector allows us to generate an estimate of the effect of the trade shock on state manufacturing output. The countries and industries driving these results also are highlighted.

The Asian crisis affected the U.S. economy through several channels, most notably, a direct trade effect and indirect commodity price and interest

[1] In this article, East Asia is defined as China, Hong Kong, Indonesia, Japan, Korea, Malaysia, Philippines, Singapore, Taiwan, and Thailand. These 10 countries are the ones most directly associated with the Asian crisis. Focusing attention on a subset of the seven most affected countries does not change the nature of our results.

[2] A microeconomic analysis of the Asian crisis focusing on firm-level effects can be found in the article by Emmons and Schmid in this issue of the *Review*.

[3] For examples of analyses focusing on specific states, see Valletta (1998) for a study of California, Oregon, and Washington; Duca et al. (1998) for a study of Texas; and Gazel and Lamb (1998) for a study of Colorado, Kansas, Missouri, Nebraska, New Mexico, Oklahoma, and Wyoming.

Table 1

State Exports of Manufactured Goods (1997)

	Total (millions of 1996 dollars)		Per Capita (1996 dollars)	
State	**East Asia (Rank)**	**All Countries (Rank)**	**East Asia (Rank)**	**All Countries (Rank)**
California	$49,333 (1)	$104,726 (1)	$1,533 (3)	$3,254 (5)
Texas	15,327 (2)	80,178 (2)	791 (9)	4,136 (3)
Washington	14,418 (3)	31,996 (5)	2,568 (1)	5,699 (2)
New York	7,697 (4)	38,881 (3)	424 (24)	2,143 (17)
Arizona	6,329 (5)	14,466 (13)	1,390 (4)	3,177 (6)
Illinois	6,259 (6)	28,570 (6)	522 (18)	2,383 (11)
Massachusetts	5,385 (7)	17,454 (9)	881 (7)	2,855 (8)
Ohio	4,655 (8)	26,286 (8)	416 (25)	2,348 (12)
Oregon	4,086 (9)	8,294 (21)	1,260 (5)	2,557 (9)
Pennsylvania	3,754 (10)	17,392 (10)	313 (35)	1,448 (30)
North Carolina	3,218 (11)	16,789 (11)	433 (22)	2,259 (13)
Minnesota	3,188 (12)	9,644 (20)	680 (10)	2,057 (19)
Georgia	3,024 (13)	13,904 (14)	404 (26)	1,856 (23)
Florida	2,760 (14)	26,438 (7)	188 (44)	1,801 (26)
New Jersey	2,759 (15)	15,833 (12)	342 (33)	1,965 (21)
Michigan	2,754 (16)	33,886 (4)	282 (37)	3,465 (4)
Virginia	2,555 (17)	12,322 (16)	379 (29)	1,829 (24)
Louisiana	2,545 (18)	10,996 (17)	585 (12)	2,526 (10)
Colorado	2,066 (19)	5,470 (28)	531 (16)	1,405 (34)
Indiana	2,036 (20)	12,979 (15)	347 (32)	2,213 (15)
Wisconsin	2,018 (21)	10,587 (18)	388 (27)	2,035 (20)
Tennessee	1,797 (22)	9,783 (19)	334 (34)	1,821 (25)
Connecticut	1,715 (23)	7,325 (24)	525 (17)	2,242 (14)
New Mexico	1,518 (24)	1,866 (38)	881 (8)	1,082 (41)
Kentucky	1,419 (25)	8,276 (23)	363 (31)	2,116 (18)
South Carolina	1,375 (26)	8,291 (22)	363 (30)	2,189 (16)
Kansas	1,345 (27)	4,490 (30)	517 (19)	1,726 (27)
Maryland	1,334 (28)	5,774 (27)	262 (40)	1,133 (39)
Alabama	1,247 (29)	6,201 (26)	288 (36)	1,435 (32)
Iowa	1,097 (30)	5,344 (29)	384 (28)	1,872 (22)
Utah	984 (31)	3,204 (32)	477 (20)	1,552 (29)
Missouri	977 (32)	7,081 (25)	181 (46)	1,309 (36)
Vermont	960 (33)	4,129 (31)	1,632 (2)	7,015 (1)
Nebraska	957 (34)	2,110 (37)	577 (13)	1,274 (37)
Idaho	800 (35)	1,744 (39)	661 (11)	1,443 (31)
Alaska	760 (36)	1,024 (44)	1,246 (6)	1,680 (28)
Maine	707 (37)	1,704 (40)	569 (14)	1,372 (35)
Oklahoma	705 (38)	2,968 (33)	212 (43)	894 (45)
Arkansas	685 (39)	2,484 (35)	271 (39)	985 (43)
Mississippi	461 (40)	2,642 (34)	169 (48)	967 (44)
West Virginia	459 (41)	1,563 (42)	253 (41)	861 (47)
New Hampshire	328 (42)	1,665 (41)	280 (38)	1,420 (33)
Delaware	316 (43)	2,233 (36)	430 (23)	3,037 (7)
Wyoming	269 (44)	576 (47)	561 (15)	1,200 (38)
District of Columbia	244 (45)	468 (49)	460 (21)	884 (46)
Rhode Island	240 (46)	1,034 (43)	243 (42)	1,047 (42)
Hawaii	219 (47)	288 (51)	184 (45)	242 (51)
Nevada	164 (48)	936 (45)	98 (49)	557 (49)
South Dakota	126 (49)	504 (48)	171 (47)	683 (48)
Montana	70 (50)	393 (50)	80 (50)	448 (50)
North Dakota	35 (51)	713 (46)	54 (51)	1,113 (40)

rate effects.[4] While most of this article focuses on the effect of the trade channel on state manufacturing sectors, we attempt to provide some understanding of the relative importance of this effect. To do so, we examine both the strength of the trade shock and a measure of the indirect effects (oil prices) as determinants of state manufacturing employment growth. This analysis provides a rough estimate of the extent to which the indirect effects of the Asian crisis may have offset the direct effects of the crisis.

STATE MANUFACTURING EXPORTS TO EAST ASIA

Across states, the levels of manufacturing exports to all countries and manufacturing exports to East Asia vary substantially.[5] During 1997, California was the leading state with manufacturing exports totaling $104.7 billion, while Hawaii had the smallest amount of manufacturing exports—$0.29 billion as shown in Table 1.[6] California also was the state with the largest amount of manufacturing exports to East Asia—$49.3 billion. North Dakota had the smallest amount of manufacturing exports to East Asia—$0.04 billion.

Because larger states tend to have higher levels of manufacturing exports, we also have presented manufacturing exports on a per-capita basis in Table 1. On this basis, Vermont, Washington, Texas, Michigan, and California were the five leading states. All of these five states, except Vermont, also are among the top five exporting states on a gross-dollar basis. Substantial differences exist on a per-capita basis between the 10 largest exporters and the 10 smallest exporters. The average per-capita value of the 10 largest exporting states was $3,772, while the average of the 10 smallest exporters was $757.

Turning to per-capita exports to East Asia, Washington, Vermont, California, Arizona, and Oregon were the five leading states. Washington, California, and Arizona were among the top five states in exports to East Asia on a gross-dollar basis. Oregon was in the top 10, but Vermont was much further down the list. Once again, substantial differences exist between the 10 largest and 10 smallest exporters. The average per-capita value of the 10 largest exporters to East Asia was $1,286, while the average of the 10 smallest exporters was $158.

In light of the national decline in exports to East Asia from 1997 to 1998, it is not surprising that the exports of most states declined, as shown in Table 2. Ten of the 51 states (the District of Columbia is treated as a state in this paper) experienced a rise in exports to East Asia, while the remaining 41 states experienced declines. Thirty-one of these latter states had double-digit percentage declines in exports to East Asia. Eight states—Alaska, Arizona, the District of Columbia, Maryland, Montana, Nevada, North Dakota, and Vermont—saw their exports fall by more than 30 percent.

Examining state worldwide manufacturing exports provides additional perspective on the Asian trade shock. While only 10 of 51 states experienced a rise in exports of manufactured goods to East Asia, 26 states saw their worldwide exports of manufactured goods rise. Nevertheless, there is a high positive correlation between the changes in a state's manufacturing exports to East Asia and the changes in its worldwide manufacturing exports.[7] All 10 states with rising manufacturing exports to East Asia were among the 26 states whose worldwide manufacturing exports rose, while 25 of the 41 states whose manufacturing exports to East Asia fell, also saw a decline in worldwide manufacturing exports.

THE ASIAN CRISIS TRADE SHOCK AND STATE MANUFACTURING OUTPUT

The effect of the Asian trade shock on an individual state can be separated into two factors: 1) the change in manufacturing exports to East Asia and 2) the importance of those exports to the state's economy. As shown in Table 2, 10 states increased their manufacturing exports to East Asia during 1998. These states, especially the five states with double-digit increases, appear to have been immune to the Asian crisis.[8] Turning to the states whose manufacturing exports to East Asia fell during 1998, the declines ranged from 0.7 percent in Arkansas to 56.7 percent in Montana. Despite being suggestive, these data are not sufficient to conclude that the economy

[4] See Noland et al. (1998) and McKibbin (1998) for discussions of general equilibrium models and estimates indicating negligible macroeconomic effects.

[5] See the appendix for a discussion of the state export data as well as all other data used in our study.

[6] We use 1997 as the reference year for comparing exports across states because the 1998 comparisons are affected by the Asian crisis.

[7] The correlation coefficient is 0.76.

[8] The increase in these states' exports to East Asia does not mean that they did not suffer trade effects from the Asian crisis. It is possible that the increases in exports during 1998 were below what would have occurred in the absence of the crisis. To examine such a hypothesis one needs to know the trend in exports to East Asia from these states. Unfortunately, the MISER data prior to 1996 are not comparable with the more recent data, making it difficult to calculate a trend.

Table 2

Change in Real State Exports of Manufactured Goods

(Percent change 1997-98)

State	East Asia	All Countries
Maine	22.4	10.3
Rhode Island	22.1	7.5
New Hampshire	19.1	11.2
New Mexico	15.4	9.7
Washington	14.8	24.0
Louisiana	6.9	2.1
Oregon	5.1	4.9
Colorado	0.8	4.9
New Jersey	0.8	5.5
Florida	0.5	6.5
Arkansas	–0.7	2.1
Virginia	–4.2	0.4
Nebraska	–4.9	4.5
Tennessee	–5.2	4.0
Pennsylvania	–5.8	–0.2
Kansas	–6.9	–3.7
Illinois	–7.7	11.5
Indiana	–8.5	3.9
North Carolina	–9.6	–4.3
Idaho	–9.8	–7.0
Connecticut	–10.4	6.1
Georgia	–10.5	3.9
Minnesota	–10.9	–2.4
Massachusetts	–12.0	–2.2
Missouri	–12.7	–10.9
Delaware	–14.4	7.0
Iowa	–15.0	–0.4
Texas	–15.5	6.6
South Carolina	–16.4	4.2
New York	–16.6	–0.1
Alabama	–16.7	8.1
West Virginia	–17.4	–2.8
California	–18.5	–1.6
Oklahoma	–19.2	1.3
Utah	–19.4	5.2
Wisconsin	–19.9	–3.5
Michigan	–20.0	–8.6
Kentucky	–22.2	2.7
Wyoming	–24.8	–8.0
Hawaii	–25.6	–16.4
Ohio	–27.9	1.4
Mississippi	–28.4	–7.7
South Dakota	–29.6	–13.7
Arizona	–30.7	–16.7
Nevada	–33.0	–23.7
Vermont	–36.5	–0.7
Maryland	–40.1	–10.6
North Dakota	–40.5	–5.5
Alaska	–41.6	–33.2
District of Columbia	–46.5	–29.3
Montana	–56.7	–17.6

of Montana was affected more severely by this trade shock than the economy of Arkansas. To determine how these declines in exports affect a state's economy, one must look at the importance of these exports to output.

One clue to the importance of East Asian exports for a state is the share of that state's exports going to the region. As shown in Table 3, the geographic pattern of trade varies across states. During 1997, East Asia was the destination for 81 percent of New Mexico's manufacturing exports. Alaska and Hawaii also were highly dependent on the East Asian markets as both sent about 75 percent of their manufacturing exports to the East Asian countries. The other states bordering the Pacific Ocean—Oregon, California, and Washington—also sent a sizeable share of their manufacturing exports to East Asia. In contrast, only 5 percent of the manufacturing exports of North Dakota were shipped to East Asia. As a general statement, states in the western United States tend to ship a higher percentage of the manufacturing exports to East Asia than eastern states, as shown in Figure 1. One notable exception is the District of Columbia, which sent over 50 percent of its manufacturing exports to East Asia during 1997.

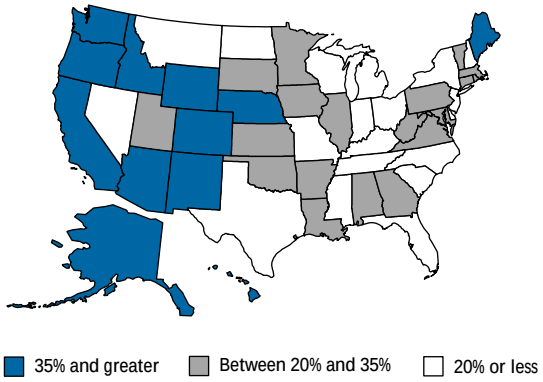
A more precise measure of how important the East Asian markets are to a state is given by a state's manufacturing exports to the region as a share of its manufacturing output. Using this measure, Alaska's manufacturing sector was the most dependent on East Asia—one-quarter of its manufactured shipments were sent to this region.[9] In contrast, less than 1 percent of manufactured shipments from firms in North Dakota were sent to East Asia. Thus, if exports to East Asia fell by the same amount in Alaska and North Dakota, the effect on the Alaskan economy would be 25 times greater. The data in Table 4, as in Table 3, indicate the relative dependence of western states on the East Asian markets.

Multiplying a state's change in manufacturing exports to East Asia (Table 2) by the share of those exports in manufacturing shipments (Table 4) provides an estimate of the trade effect of the Asian crisis on a state's economy as shown in Table 5. More precisely, the data in Table 5 indicate the

[9] A better measure of the effect of the change in exports on a state's economy would be the share of exports in a state's gross product. A state's exports are measured, however, by the total dollar value of the shipments while gross state product is based on value added.

Figure 1

The Importance of East Asia for a State's Manufacturing Exports 1997



contribution of manufactured exports to East Asia to a state's manufacturing sector growth rate during 1998. For example, the 26-percent decline in Hawaii's exports to East Asia reduced the growth rate of manufacturing output in that state by 1.5 percentage points. While these estimates indicate that the East Asian trade shock reduced the growth rate of manufacturing output by less than 0.3 percentage points in most states, they also indicate the diversity of effects across states even if one considers only those states whose exports to East Asia fell.

The data in Table 5 do not measure the overall change in manufacturing output in the states during 1998. Rather, these data denote the contribution of exports to the growth rate of manufacturing output. For example, the 19-percent decline in Utah's exports to East Asia (Table 2) reduced the growth rate of manufacturing output in that state by 1 percentage point (Table 5). It is possible, however, that this decrease in exports to East Asia was offset by either an increase in exports to other foreign markets or an increase in domestic sales. Overall, Utah's manufacturing sector may have experienced no decline in growth, depending on the strength of demand in these other markets. The data in Table 2 indicate that for most states a decline in exports to East Asia was not offset by a rise in exports to other regions. Whether domestic demand was strong enough to overcome the trade effect will not be known until the 1998 shipments data are released in late 2000.

Table 3

East Asia's Share of Real State Exports of Manufactured Goods (1997)

State	Percent Share
New Mexico	81.4
Hawaii	76.0
Alaska	74.2
District of Columbia	52.1
Oregon	49.3
California	47.1
Wyoming	46.8
Idaho	45.9
Nebraska	45.3
Washington	45.1
Arizona	43.8
Maine	41.5
Colorado	37.8
Minnesota	33.1
Massachusetts	30.9
Utah	30.7
Kansas	30.0
West Virginia	29.4
Arkansas	27.6
South Dakota	25.0
Oklahoma	23.7
Connecticut	23.4
Vermont	23.3
Rhode Island	23.2
Louisiana	23.1
Maryland	23.1
Illinois	21.9
Georgia	21.7
Pennsylvania	21.6
Virginia	20.7
Iowa	20.5
Alabama	20.1
New York	19.8
New Hampshire	19.7
North Carolina	19.2
Texas	19.1
Wisconsin	19.1
Tennessee	18.4
Montana	17.8
Ohio	17.7
Nevada	17.5
Mississippi	17.4
New Jersey	17.4
Kentucky	17.1
South Carolina	16.6
Indiana	15.7
Delaware	14.2
Missouri	13.8
Florida	10.4
Michigan	8.1
North Dakota	4.9

Table 4

Manufacturing Exports to East Asia as a Share of Manufactured Shipments (1996)

State	Export-Shipment Ratio
Alaska	25.0
Washington	17.3
California	13.4
Arizona	12.7
Oregon	8.8
Wyoming	8.4
Hawaii	5.9
Utah	5.6
Massachusetts	5.3
Colorado	4.9
Texas	4.6
Idaho	4.6
New York	4.3
New Mexico	4.2
Vermont	4.1
Maryland	3.9
Minnesota	3.9
Nebraska	3.8
Illinois	3.3
District of Columbia	3.3
Florida	3.2
Connecticut	3.1
Louisiana	2.9
Virginia	2.9
New Jersey	2.8
Kansas	2.7
Maine	2.7
Georgia	2.6
West Virginia	2.6
South Carolina	2.2
Nevada	2.2
Ohio	2.1
Rhode Island	2.0
Alabama	2.0
Pennsylvania	1.9
Iowa	1.9
New Hampshire	1.9
North Carolina	1.8
Oklahoma	1.8
Tennessee	1.8
Wisconsin	1.6
Delaware	1.5
Indiana	1.5
Kentucky	1.5
Arkansas	1.4
Michigan	1.4
Montana	1.2
South Dakota	1.0
Missouri	0.9
Mississippi	0.9
North Dakota	0.6

A CLOSER LOOK AT THE STATE EFFECTS

Country Detail

Examining the change in state manufacturing exports to each country in East Asia, rather than to the region as a whole, may provide some insight into the different effects across states. At first glance, differences in the states' trading partners do not appear to explain the variations in effects across states. First, exports declined to nearly all countries in the region, as underscored by the data in Table 6. Only three states—Maine, New Hampshire, and Washington—experienced declines in exports to fewer than half of the East Asian countries. Not surprisingly, these three states were among those whose exports to East Asia rose during 1998. In contrast, most states had export declines to at least seven of the 10 East Asian countries. Four—Maryland, Montana, North Carolina, and Ohio—had declines in exports to all 10 countries. As Table 2 shows, these states, except for North Carolina, had export declines of 25 percent or more to the region.

Second, across states, the major trading partners in East Asia do not vary substantially. Japan is the most important export destination in East Asia for most states, and South Korea is an important export market for many. As a result, declines in exports to these two countries typically accounted for a large share of a state's overall decline in exports to East Asia in 1998.

Nevertheless, for some states, trading partners mattered. The economy of Thailand, for example, suffered one of the most severe contractions in the region with output falling by 8 percent during 1998. Thailand was not an important export market for most states, but it was for the District of Columbia and Minnesota. In 1997, 48 percent of D.C.'s exports of manufactured goods to East Asia went to Thailand. In 1998, D.C.'s exports to Thailand fell by 97 percent, accounting for nearly all of the 47 percent decline in its exports of manufactured goods to East Asia. Meanwhile, Thailand accounted for 13 percent of Minnesota's manufacturing exports to East Asia during 1997. These exports fell by half in 1998, accounting for 60 percent of Minnesota's overall decline in manufacturing exports to East Asia. On the other hand, China's economy remained relatively robust during the crisis. China is not a major export destination for many states; however, Louisiana sent over 20 percent of its 1997 East Asian exports to China. A rise in these exports during 1998 accounted for the

overall increase in Louisiana's exports to the region.

Industry Detail

Disaggregating the manufacturing data from the one-digit SIC level to the two-digit SIC level also provides some insight into the different experiences of the states during 1998. The chemical and allied products (SIC 28), industrial machinery and equipment (SIC 35), and electronic and electrical products (SIC 36) industries represented the highest shares of manufacturing exports to East Asia across the broadest range of states. Each of these industries accounted for more than 10 percent of manufacturing exports to East Asia in more than half of the states. Food and kindred products (SIC 20), transportation equipment (SIC 37), and instruments and related products (SIC 38) also were important industries for a number of states exporting to East Asia. The importance of individual manufacturing industries for exports varies across states, more so than the destination of these exports, as discussed above.

The manufacturing exports of some states are concentrated in a single industry. The electronic and electrical products industry, for example, accounted for 96 percent of New Mexico's manufacturing exports to East Asia during 1997, while chemicals and allied products accounted for 99 percent of Wyoming's exports to the region. A sharp drop in Wyoming's chemical exports to East Asia, in conjunction with the relative importance of these markets for the state industry, was responsible for most of the negative effect on the manufacturing sector noted in Table 5. Wyoming's experience is in contrast with that of Washington. During 1997, 64 percent of Washington's manufactured exports to East Asia were transportation equipment, mostly aircraft. Sales of aircraft to East Asia, particularly China and Taiwan, increased sharply during 1998. These increases primarily are responsible for the increase in Washington's manufacturing exports to East Asia.[10]

The manufacturing exports of other states were more diversified. For example, none of the 20 industries accounted for more than 20 percent of the exports to East Asia from either Missouri, North Car-

[10]It is likely that the increase in aircraft exports accounts for most of the positive export boost in Washington's manufacturing sector during 1998, listed in Table 5. Manufacturing shipments for Washington's transportation industry are not disclosed by the Department of Commerce, however, to prevent disclosure of data relevant to one company.

Table 5

The Effect of the Trade Shock on Manufacturing Output Growth

State	Percent
Washington	2.56
New Mexico	0.65
Maine	0.60
Oregon	0.45
Rhode Island	0.44
New Hampshire	0.36
Louisiana	0.20
Colorado	0.04
New Jersey	0.02
Florida	0.02
Arkansas	–0.01
Tennessee	–0.09
Pennsylvania	–0.11
Missouri	–0.12
Virginia	–0.12
Indiana	–0.13
North Carolina	–0.18
Kansas	–0.19
Nebraska	–0.19
Delaware	–0.22
North Dakota	–0.25
Illinois	–0.26
Mississippi	–0.26
Michigan	–0.27
Georgia	–0.28
Iowa	–0.28
South Dakota	–0.31
Connecticut	–0.32
Alabama	–0.33
Wisconsin	–0.33
Kentucky	–0.33
Oklahoma	–0.35
South Carolina	–0.37
Minnesota	–0.42
Idaho	–0.45
West Virginia	–0.45
Ohio	–0.59
Massachusetts	–0.63
Montana	–0.71
New York	–0.71
Nevada	–0.72
Texas	–0.72
Utah	–1.08
Vermont	–1.51
Hawaii	–1.51
District of Columbia	–1.54
Maryland	–1.58
Wyoming	–2.07
California	–2.48
Arizona	–3.91
Alaska	–10.40
Unweighted average	**–0.62**

Table 6

Declining Exports to East Asia on a Country Basis (1998)

State	Number
Alaska	8
Alabama	7
Arkansas	7
Arizona	7
California	9
Colorado	6
Connecticut	8
District of Columbia	8
Delaware	7
Florida	6
Georgia	9
Hawaii	7
Iowa	8
Idaho	7
Illinois	8
Indiana	9
Kansas	8
Kentucky	7
Louisiana	7
Massachusetts	7
Maryland	10
Maine	3
Michigan	9
Minnesota	5
Missouri	6
Mississippi	9
Montana	10
North Carolina	10
North Dakota	9
Nebraska	5
New Hampshire	3
New Jersey	7
New Mexico	7
Nevada	8
New York	6
Ohio	10
Oklahoma	7
Oregon	6
Pennsylvania	7
Rhode Island	5
South Carolina	9
South Dakota	6
Tennessee	7
Texas	8
Utah	7
Virginia	7
Vermont	6
Washington	4
Wisconsin	9
West Virginia	9
Wyoming	8

olina, or Nevada. Exports declined across a broad range of the industries in all of these states.

THE ASIAN TRADE SHOCK AND STATE MANUFACTURING EMPLOYMENT

The results in Table 5 indicate that the manufacturing sector in some states was subjected to large, negative shocks, while many other states were affected only slightly. We use regression analysis to assess the importance of these trade effects on state manufacturing employment growth. Of course, the estimated trade effects were not the only influence on employment growth across states during 1998. Consequently, additional variables are required for the regression analysis. We use two additional variables, one of which is related to the Asian crisis. Commodity prices fell during 1998 partly as a result of decreased demand in East Asia. Perhaps most important was the fall in the price of oil. Those states with a high concentration of manufacturing industries that use petroleum products extensively as an input would benefit relative to states that produce petroleum and its related products. Employment trends across states, which reflect the interaction of various other economic factors affecting manufacturing employment, also are likely to be important. For example, states where manufacturing employment growth has been falling recently may be those states with industries shedding employment to remain competitive. Hence, employment changes are occurring regardless of the strength of the Asian economies.

In sum, we can think of state-level manufacturing employment growth during 1998 as being determined by previous employment growth and two shocks: the Asian trade shock, which was unfavorable to employment growth in most states; and the oil price shock, which was favorable to employment growth in most states. These relationships can be summarized by the following equation:

$$(1) \qquad megt_i = \beta_0 + \beta_1 * trade_i + \beta_2 * oil_i + \beta_3 * megtrend_i + \epsilon_i$$

where the subscript *i* refers to individual states, *megt* is the manufacturing employment growth rate in each state, *trade* is the *negative* of the estimated trade effect on the manufacturing sector in each state (given in Table 5), *oil* is an estimate of the differential effect of a decline in oil prices on each state (see the appendix), *megtrend* is the average annual growth rate of manufacturing employment from 1994 to

1997, and ϵ is an error term. The betas indicate the effect of each of these variables on manufacturing employment growth.

We expect $\beta_1 < 0$, and β_2 and $\beta_3 > 0$. First, a larger Asian trade shock should result in a larger decline in manufacturing employment growth. Specifically, states such as Arizona and Alaska should experience larger declines in manufacturing employment growth than Arkansas or Tennessee. Second, a drop in the price of oil should raise manufacturing employment growth in energy-importing states. In other words, states with industry compositions weighted toward users rather than producers of energy should experience larger increases in employment growth than others. Finally, states that have experienced recent increases in manufacturing employment are likely to continue to do so.

Our estimation of equation 1 produced the following results:

$$(2)\quad \underset{}{megt} = \underset{(0.99)}{0.23} - \underset{(-0.24)}{0.03} * trade + \underset{(3.03)}{1.93} * \text{oil} + \underset{(6.23)}{0.57} * megtrend,$$

where the t-statistics are given in parentheses.[11] These results indicate that the Asian trade shock had a negative, but statistically insignificant, effect on employment growth across states. That is, based on the regression analysis, the Asian trade shock was not a factor driving differences in state-level manufacturing employment growth during 1998. Meanwhile, we find the oil price shock to have a positive, statistically significant effect on manufacturing employment growth. Thus, statistically speaking, the differential effect of the Asian crisis on state employment may be more pronounced through its effect on oil prices than through trade flows. Finally, prior manufacturing employment growth was found to have a positive, statistically significant effect on current (1998) manufacturing employment growth.[12]

Our results leave an important question unanswered: Why didn't the trade effect have a noticeable effect on manufacturing employment across states? One possibility is that the differential shocks were not large enough to generate statistically significant differences in employment growth. Despite much variation across states in their percentage declines of exports to East Asia, and the importance of East Asia as an export destination, there is little difference across many states in the estimated trade shock. More than two-thirds of the estimates for the individual states are between 0.04 percent and –0.72 percent, a range of only 0.76 percentage points. Another possibility is that our measure of the trade shock, because it relies on shipments data (both for exports and output) rather than value-added data, is deficient. Employment changes in a state are likely to be related to a state's value-added changes; however, the shipments data might be a poor measure of the production that occurred in a specific state, and thus, are not directly related to employment.

CONCLUSION

The Asian crisis resulted in a decline in most states' exports to East Asia, but the severity of the decline varied across states. An assessment of the importance of the decline in exports for a state's economy depends on the extent of the decline and the importance of the East Asian markets for the state. In general, the western states were more dependent on the East Asian markets, and hence, were the hardest hit by the trade shock. Of the states in which the decline in exports to East Asia lowered the growth rate of manufacturing output by more than 1 percent, two-thirds were western states. Some western states, most notably Washington, however, were among those states whose exports to East Asia rose despite the crisis.

Using these estimates of the trade shock's effect on each state's manufacturing sector, we tested the statistical relevance of the trade shock to explain changes in a state's manufacturing employment during 1998. We found that the trade shock did not explain differences in the employment experiences of states during 1998. A factor that was statistically relevant was the effect of the change in the price of oil on a state's economy. Thus, taking our results at face value, the differential state employment effects that resulted from the trade changes of the Asian crisis mattered little. A stronger case can be made that the oil price declines during late 1997 and 1998, some portion of which can be attributed to the Asian crisis, were more important than the trade effects in affecting manufacturing employment at the state level.

[11] The $\bar{R}^2$ for the equation is .45.

[12] To investigate whether the Asian crisis affected state employment in specific industries, we estimated equation 1 for manufacturing industries at the two-digit level. Unfortunately, data limitations affected this effort. Only 14 of the 20 SIC industries had complete data for at least half of the states. For 11 of the 14 industries, the results were similar to those for manufacturing as a whole—the Asian trade effect, even though exhibiting the anticipated negative sign, was statistically insignificant. Results for the employment effect of the oil price shock also were not strong as the effect was statistically insignificant in 12 of the 14 industries.

REFERENCES

Brown, Stephen P.A., and Mine K. Yücel. "Energy Prices and State Economic Performance," Federal Reserve Bank of Dallas *Economic Review* (Second Quarter 1995), pp. 13-23.

Coughlin, Cletus C., and Thomas B. Mandelbaum. "Measuring State Exports: Is There a Better Way?" this *Review* (July/August 1991), pp. 65-79.

Cronovich, Ron, and Ricardo Gazel. "How Reliable Are the MISER Foreign Trade Data?" unpublished manuscript, May 1999.

Duca, John V., David M. Gould, and Lori L. Taylor. "What Does the Asian Crisis Mean for the U.S. Economy?" Federal Reserve Bank of Dallas *Southwest Economy* (March/April 1998), pp. 1-6.

Emmons, William R., and Frank A. Schmid. "The Asian Crisis and the Exposure of Large U.S. Firms," this *Review* (January/February 2000). pp. 15-34.

Gazel, Ricardo C., and Russell L. Lamb. "Will the Tenth District Catch the Asian Flu?" Federal Reserve Bank of Kansas City *Economic Review* (Second Quarter 1998), pp. 9-26.

McKibbin, Warwick. *The Crisis in Asia: An Empirical Assessment*, Brookings Discussion Paper in International Economics No. 136, 1998.

Noland, Marcus, Li-Gang Liu, Sherman Robinson, and Zhi Wang. *Global Economic Effects of the Asian Currency Devaluations*, Institute for International Economics, 1998.

Pollard, Patricia S., and Cletus C. Coughlin. "Going Down: The Asian Crisis and U.S. Exports," this *Review* (March/April 1999), pp. 33-45.

U.S. Department of Commerce and U.S. Department of Treasury. *Treasury and Commerce Release Analysis Showing Impact of Asian Crisis on Individual States*, March 24, 1998.

Valletta, Rob. "East Asia's Impact on Twelfth District Exports," Federal Reserve Bank of San Francisco *Economic Letter* (November 20, 1998).

Appendix-Data Sources

State Export Data

The data on state manufacturing exports used in this study are produced by the Massachusetts Institute for Social and Economic Research (MISER) at the University of Massachusetts. These data are export shipments by the state of origin of movement. The source of the data is the Shipper's Export Declaration (SED). This document identifies "the state where the product began its journey to point of export." The Census Bureau collects these data, which are adjusted by MISER to fill in missing industry and state information.

The MISER export data have their weaknesses. The identified export state may not be the state of manufacture, but rather the state of a broker (or wholesaler) or the state where a number of shipments were consolidated. This problem is more pronounced for exports of agricultural commodities than manufactured goods. Hence, our study focuses on manufactured goods.

An alternative to the export data based on the origin of movement is one based on exporter location. These data, which had been available since 1993, also are based on the Shipper's Export Declaration. Compared to the origin of movement series, the identified state of the exporter in these data more likely reflects the state of a broker or wholesaler, or the headquarters of companies, rather than the state of manufacture. A potentially better source of state export data for identifying the state of production has been produced by the U.S. Census as part of the Annual Survey of Manufacturers; however, *Exports from Manufacturing Establishments* was discontinued after data for 1992 were published.

The data we used, despite their limitations, are regarded as the best available on state exports.[1] They are available quarterly with a three-month lag. In addition to the state information, the data contain information on one-and two-digit industry (SIC) code, destination country, and dollar value and weight by method of transportation.[2]

[1] Additional discussion of the various issues involving state export data can be found in Coughlin and Mandelbaum (1991) and Cronovich and Gazel (1999).

[2] Information on this database is available at < www.umass.edu/miser/axes > .

Export Price Data

Real exports were calculated by deflating the 1997 export data by the change in the price of exports between 1996 and 1997. The 1998 exports were deflated by the price change between 1996 and 1998. Export price data are available from the Bureau of Labor Statistics. These data are not available by SIC code. Thus, we started with an export price index that groups the data based on the Standard International Trade Classification (SITC) system and matched these industries with the appropriate SIC codes. When multiple SITC codes fit one SIC category, a weighted average of the price indices for those categories was constructed to arrive at the price index on an SIC basis. For more details see Pollard and Coughlin (1999).

Per-Capita Exports

Per-capita exports were calculated by dividing real exports for each state by the population of the state. The population data are available from the Census Bureau.

Manufacturing Shipments Data

To measure manufacturing output at the state level we use the value of industry shipments. These data are from the Annual Survey of Manufacturing conducted by the Census Bureau. The latest available data were for 1996.

Employment Data

The state level employment data came from the payroll employment survey conducted by the Bureau of Labor Statistics. We used the average annual employment data for 1997 and 1998 to calculate the growth rate in employment during 1998. The trend employment growth rate is calculated using the average annual growth rate in state employment between 1994 and 1997. Data at both the one-and two-digit industry (SIC) code were used in the regression analysis.

Oil Price Effect Data

The oil price effect data are based on estimates by Brown and Yücel (1995) of the effect of a 10-percent change in the price of oil on a state's nonagricultural employment. Declining energy prices should retard economic activity in energy-exporting states (i.e., those that produce more energy than they consume), and stimulate economic activity in energy-importing states. The paper contains estimates for 1992 and 2000. The results reported in this paper use the estimates for the year 2000, but the results are similar using the 1992 estimates.

William R. Emmons is a economist and Frank A. Schmid is a senior economist at the Federal Reserve Bank of St. Louis. Judith Hazen, Robert Webb, and Marcela Williams provided research assistance.

The Asian Crisis and the Exposure of Large U.S. Firms

William R. Emmons and Frank A. Schmid

The financial and economic crises that ravaged Thailand, Indonesia, South Korea, Malaysia, and other Asian countries during 1997 and 1998 triggered one of the most abrupt and severe economic slowdowns seen anywhere in the world during recent decades. Financial-market volatility increased around the globe soon after the Thai devaluation of July 1997, reaching its high point in October 1998. Many countries were not hit directly by this financial crisis; nonetheless, they felt significant repercussions. Worldwide economic growth slowed, risk premiums in debt markets increased, stock markets became more volatile, and confidence indicators slumped in many countries (see *Economic Report of the President*, 1999, pp. 227-51, for an extensive discussion of the Asian crisis).

We examine how the Asian financial crisis affected the sensitivity of large U.S. firms to U.S. stock-market risk—that is, whether the economic situation in Asia is related to changes in firms' "betas." Following corporate finance theory, we define stock-market exposure as the extent to which a firm's stock returns are correlated with overall market returns. Exposure is summarized by the firm's estimated beta, which, according to the Capital Asset Pricing Model (CAPM), is simply the coefficient estimated in a regression of the firm's excess returns—i.e., dividends plus price change less the risk-free return—on market excess returns over some specified period.[1] If a firm's beta rises, investors demand higher excess returns for holding its stock. This raises the firm's cost of equity capital. Thus, firms that saw their betas rise as a result of the crisis would face higher equity financing costs.

If the Asian crisis mattered at all for U.S. firms' stock-market risk, we would expect the largest effects at firms with the highest relative sales exposure to Asia. Similarly, we would speculate that the betas of firms with relatively low sales exposure to the Asian region would remain unchanged or decrease. Betas are measures of a firm's return sensitivity *relative* to the market, so some betas must go down if others go up. In addition, we would expect firm leverage (debt as a percentage of assets) to amplify the effects of the Asian crisis on firms' CAPM betas. This is because contractually fixed payments owed to debtholders do not change even if underlying cash flows are reduced. This is in contrast to the situation of equity investors, who hold the residual cash flow rights. The more highly leveraged a firm is, the larger the income variability to which the equity holders are exposed, everything else being equal. This income variability is related directly to the amount of total assets the firm controls or its sales, not to its equity base.

Our analysis begins by identifying the S&P 100 firms that reported detailed regional breakdowns of sales for 1996-98.[2] We then estimate a model of each firm's weekly excess stock returns for the period 1997-98. We use the weekly excess return on the S&P 500 as the relevant market factor because we want a broad measure for the market. Our model is based on the Capital Asset Pricing Model but extends it to allow for changes in betas, perhaps related to developments in Asia.[3] Intuitively, betas may change in response to the Asian crisis or other shocks to the fundamentals of firms that are not shared to the same degree by the market as a whole.

We run a second-stage regression using results of the first-stage asset-pricing model as the dependent variable. We analyze the sensitivity of the firms' betas to the extent of their sales exposures to Asia, where sales are weighted by the ratio of long-term debt to total assets. We find that leverage-weighted sales exposure to Asia exerts a significant positive effect on a firm's CAPM beta.

---

[1] See Brealey and Myers (1996, pp. 160-64).

[2] We attempted but were not able to obtain a breakdown of sales by world regions for all of the S&P 100 firms. Only 39 firms had sufficiently detailed information to be included in our sample.

[3] For a textbook presentation of the CAPM, see Brealey and Myers (1996, pp. 173-203). Jagannathan and Wang (1996) and others have developed the idea of a "conditional CAPM," that is, a period-by-period CAPM that allows betas to vary over time.

Firms that had high sales exposure to Asia became more sensitive to movements in the S&P 500 while firms with low sales exposure to the Asian region became less sensitive.

THE ASIAN CRISIS AND THE ROLE OF INCREASED BUSINESS RISK

Following several decades of rapid economic growth and increasing integration into global capital markets, the economies of a number of Asian countries suffered abrupt and severe contractions during 1997 and 1998 (*Economic Report of the President*, 1999, pp. 227-51). The proximate cause of the crisis was a failed devaluation in Thailand, but recent analyses of the period have pointed to structural and especially financial-sector weaknesses in many of the affected countries. At the same time, a very large inflow of foreign capital during the early and mid-1990s, and the subsequent rapid withdrawal of many foreign investors during 1997 and 1998, exacerbated the adjustment process.

The reverberations of the Asian crisis on the world economy have been multifaceted. World economic growth slowed as the shortfall in demand from the Asian region caused both a severe regional recession and a deterioration in the trade balances of important trading partners such as the United States (Pollard and Coughlin, 1999). Commodity prices weakened, export competition increased in many sectors, and interest rates fell in the world's major economies. Industrial production slowed in many countries and corporate profits decelerated or declined.

Despite these disruptions, the U.S. economy grew strongly throughout the 1997-98 period. This experience has led some observers to conclude that the growth-enhancing consequences of the crisis for the United States—primarily lower interest rates and lower commodity prices—were simply more powerful than the growth-reducing factors, which included reduced demand for U.S. exports and financial losses suffered by lenders and investors in the region.

Another impact of the financial crisis was an increase in the observed volatility of financial markets and capital flows around the world. Figure 1 shows the sustained increase in stock-market volatility that occurred in the United States throughout 1997-98. The higher implied volatility of the S&P 100 index during this period indicates that investor uncertainty about future stock-market returns had increased. In addition, capital flows to emerging markets collapsed while portfolio investments into the United States and other industrialized countries increased. These shifts in capital flows may go some way toward explaining the surge in U.S. and European stock-market price indexes during this period despite increased uncertainty about global economic growth and increased financial market volatility.

Figure 2 contrasts the divergent paths of U.S. and Asian emerging-market financial returns throughout this period. The cumulative total return during the two-year period on the S&P 100 index of large-capitalization U.S. firms was an astonishing 71 percent, nearly 60 percentage points better than the risk-free return on three-month Treasury bills. Meanwhile, the dollar-denominated total return during the two years on the FT/S&P Actuaries Pacific Stock-Market Index (excluding Japan) was about negative 42 percent. An investment in money-market instruments issued by Asian emerging-markets borrowers that was continually rolled over during 1997 and 1998 would have earned about a 2 percent total return in dollar terms, according to the J.P. Morgan Emerging Local Markets Index Plus (Asia).

The surge in U.S. stock prices is even more remarkable when one considers that after-tax earnings per share of the S&P 500 companies actually fell during 1997 and 1998, and that corporate bond yield spreads—that is, the extra yield that corporate issuers were forced to pay to borrow in comparison to U.S. Treasury yields—rose considerably (see Figure 3). Although alternative explanations may exist, a single risk-based explanation of these phenomena—high stock returns and volatility, declining or flat profits, and increasing bond yield spreads—is plausible. In a nutshell, the main effect of the Asian crisis, which began in the second half of 1997 and spread to Russia in August 1998, may have been to increase the perceived riskiness of corporate cash flows.

To understand how investor perceptions of increased business risk could be responsible for high stock returns and volatility, as well as increased corporate bond yield spreads—all while earnings remain flat—we need to apply the "option-theoretic interpretation of the firm."[4] The key insight of this approach is that the equity of a firm that has issued debt (e.g., bonds) is identical to a call option written on the assets of the firm, the ownership of which has been, in effect, transferred to the

[4] Black and Scholes (1973) is the original source of this interpretation of the corporate financial structure of every firm. Brealey and Myers (1996, pp. 564-66) provide a textbook discussion of these ideas.

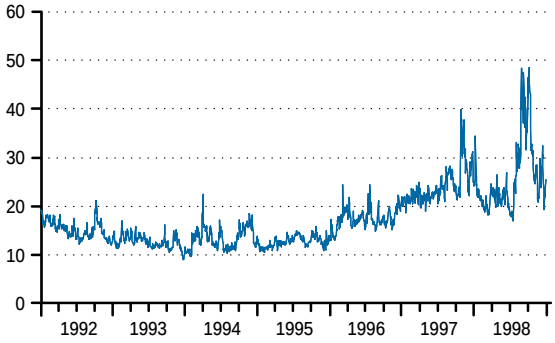
debtholders. When issuing a bond, the owners of a previously debt-free firm effectively sell the firm's assets to the bondholders and receive, in addition to the cash proceeds of the bond issue, a call option that gives them (the equityholders) the right—but not the obligation—to buy back the firm's assets by paying off the debt in full with interest. Thus, the bondholders will end up either with their money back (plus interest) or the firm itself, whichever is less valuable (because the choice is made by the equityholders). In the case of bankruptcy, the owners have decided that the firm is of less value to them as a going concern than the cash required to buy it back from its debtholders. The equityholders allow their call option to expire unexercised and "walk away" from the firm by virtue of limited liability.

How does this apply to the Asian crisis and the conjunction of high stock returns and increased corporate bond yield spreads during 1997 and 1998? One key determinant (among several) of the value of any option is the volatility of the underlying cash flows upon which the option is written. In the present application, the U.S. corporate sector generates cash flows for which corporate equityholders and corporate bondholders have claims. Precisely because the owner of an option has the right but not the obligation to exercise it, a greater dispersion of likely outcomes—i.e., higher risk—enhances the value of the option. The optionholder can capture all of the increased "upside" while ignoring all of the "downside," even if this has increased as well. This explains why stock prices could go up during the 1997-98 period even while corporate earnings were flat. Investors may have bid up stock prices because the range of future earnings estimates had increased (even if their expected level did not change). That is, the value of a call option on the assets of a firm increases when fundamental business risk increases.

Is there any other evidence that increased risk, rather than higher expected earnings, boosted stock returns during the Asian crisis? The behavior of corporate bond prices provides this evidence. The "dual" or complementary approach to corporate financial valuation within the option-theoretic interpretation of the firm can explain why corporate bond prices declined and yield spreads went up at the same time that stock prices went up. The dual approach points out the equivalence between the call-option interpretation presented above and a put-option interpretation that is relevant for corporate bondholders. In this interpretation, corporate bondholders effectively purchase an equivalent default-risk-free bond (i.e., identical in maturity and coupon rate) and underwrite a put option on the assets of the firm that is given to the equityholders. From the bondholders' perspective, the equityholders retain ownership of the firm's assets and are given the right to "put" (deliver) the assets of the firm to

Figure 1

Chicago Board Options Exchange S&P 100 Volatility Index

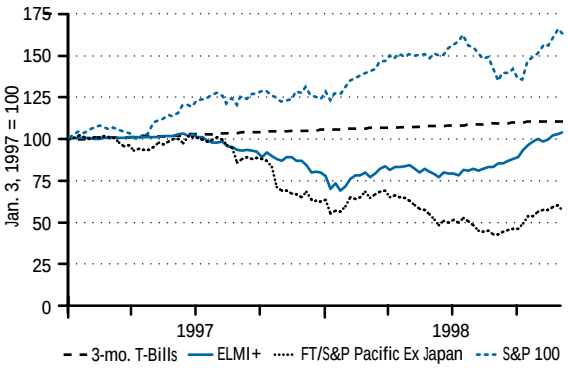


Source: Created using data from Chicago Board Options Exchange <http://www.cboe.com/tools/historical/vix.htm>.

Note: The vertical scale measures the implied annualized volatility (standard deviation) of the S&P 100 derived from a hypothetical 30-day call option written on the S&P 100 with strike price at the current day's market price, in percent.

Figure 2

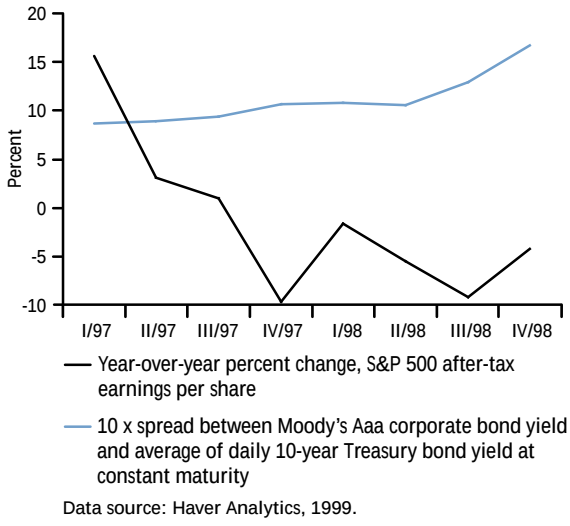
Total Return Indexes (in U.S. Dollar Terms)



Data sources: The S&P 100 return data were purchased from Standard & Poor's Compustat. Data for calculating the three-month T-Bill return are from various editions of the *Wall Street Journal*. The Asian indices were purchased from J.P. Morgan and the FT/S&P Actuaries, respectively.

Figure 3

Quarterly Corporate Earnings Growth and Corporate Bond Yield Spreads



Data source: Haver Analytics, 1999.

the bondholders in lieu of cash payment in satisfaction of the debt obligations.

Increased business risk means that the probability of default—situations when the assets are worth less than the debt they secure—has increased. Thus, increased business risk means that the put option underwritten by bondholders has become more valuable and consequently the value of the bonds has fallen. This implies that corporate bond yield spreads rise, as they did during the Asian crisis. Together, these option-theoretic interpretations of corporate securities can explain why high stock-market volatility, rising equity prices, and falling corporate bond prices—all while expected earnings have not changed—are perfectly consistent. The underlying shock was an increase in the riskiness of the corporate sector and the stock and bond price movements simply reflected a redistribution of firm value among claimholders.[5]

We would not expect that all firms' riskiness went up at the same time by the same amount. Instead, one would expect greater stock-price responses at firms with more exposure to the source of increased risk. In addition, one would expect a greater stock-price response at those firms that are more highly leveraged; this is because they are more likely to become bankrupt, all else being equal, so the option to default is more valuable. This article explores the firm-level changes in systematic risk that may have been induced by the Asian crisis. Systematic risk is measured as the sensitivity of a firm to market-wide stock-price movements. Risk is measured relative to the market as a whole. Our goal is to isolate differences in firm responses to the Asian crisis rather than to measure the overall impact on market risk.

The empirical analysis of the article uses two total-return series as alternative proxies for the evolution and severity of the Asian crisis. Although they clearly reflect different asset and risk classes, these two total-return series present a similar picture of the timing of the onset of the crisis, roughly the middle of 1997. By way of contrast, the money-market index indicates that recovery from the crisis began in early 1998, while the stock-market index did not turn upward decisively until the third quarter of 1998.

THE ECONOMETRIC MODEL: FIRST-STAGE REGRESSION

Our goal is to analyze the channel through which the Asian crisis affected the risk positions of U.S. firms during the period 1997-98.[6] We use a two-stage regression approach because it makes interpretation of the results relatively straightforward.[7] In the first-stage regression of our model, we provide evidence for the influence of the Asian crisis on the stock-market risk of the sample firms. For this purpose we set up the asset-pricing model with time-varying betas. We show that this model can be rewritten as a two-factor model, with a proxy for the Asian crisis appearing in the second factor. We provide evidence that the Asian crisis indeed affected the systematic risk of U.S. firms.

First-Stage Estimation Procedure

Our first-stage regression estimates a weekly asset-pricing model for each firm i ($i = 1,...,39$) for the period 1997-98. This model follows the CAPM but extends it in two ways to allow for time-varying β-coefficients. First, we allow the firm-specific

[5] Schmid (1999) provides a concise overview of the argument made here.

[6] All sample firms are headquartered in the United States with the exception of Northern Telecom (NORTEL) of Canada.

[7] Alternatively, we could estimate the model using a single-stage approach, as a later section describes briefly.

parameters of the model to differ between 1997 and 1998.[8] Second, we allow the CAPM beta of each firm to depend on developments in Asia, which we proxy by two alternative measures of the excess dollar return on a portfolio invested in Asian securities.

The first-stage regression model is the following:

$$R_{i,t} - R_{f,t} = \beta_{i,j}\left(1 + \lambda_{i,j} \times CRISIS_t\right) \times \left(R_{m,t} - R_{f,t}\right) + \varepsilon_{i,t} \quad (1)$$

with $j = 1997$ for $t = 1,...,52$ and $j = 1998$ for $t = 53,...,104$ (weeks running from the beginning of 1997 to the end of 1998). The dependent variable, $R_{i,t} - R_{f,t}$, is the excess log return in week t on firm i's stock in period t; $R_{f,t}$ is the risk-free rate, the one-week return on a three-month Treasury bill. $CRISIS_t$ is a variable that proxies for the state of the Asian crisis in week t; $R_{m,t} - R_{f,t}$ is the excess log return during week t on the U.S. stock-market portfolio (S&P 500); and $\varepsilon_{i,t}$ is the error term. The error term $\varepsilon_{i,t}$ measures realizations of the firm's idiosyncratic risk, that is, movements in the firm's excess return that are not explained by its comovement with the market.

The parameters of interest are $\beta_{i,j}$ and $\lambda_{i,j}$. It is important to recognize that it is the entire expression $\beta_{i,j}(1 + \lambda_{i,j} \times CRISIS_t)$ that corresponds to the market beta of firm i in the standard CAPM framework (see Brealey and Myers, 1996, pp. 179-83). In purely statistical terms, the standard CAPM beta—as well as our more complicated expression for beta—is the covariance of firm i's excess log return with the U.S. market excess log return, divided by the variance of the market excess log return. In interpreting our model, the parameter $\lambda_{i,j}$ is meant to capture the possible dependence of the CAPM beta estimated for firm i on the economic situation in Asia. We measure the current state of the Asian crisis as the weekly excess return on an Asian portfolio of securities. This is because we want to allow each firm's CAPM beta to be time-varying and to depend on investors' expectations about Asia and the particular firm's exposure to these developments.

Our time-varying extension of the CAPM model in equation 1 is nonlinear in the parameters. For ease of estimation we remove the nonlinearity by defining a new parameter, $\delta_{i,j} = \beta_{i,j}\lambda_{i,j}$, and rewriting equation 1 as follows:

$$R_{i,t} - R_{f,t} = \beta_{i,j}\left(R_{m,t} - R_{f,t}\right) + \delta_{i,j} \times CRISIS_t \times \left(R_{m,t} - R_{f,t}\right) + \varepsilon_{i,t} \quad (2)$$

again with $j = 1997$ for $t = 1,...,52$ and $j = 1998$ for $t = 53,...,104$. The firm's exposure to market movements, its CAPM or market beta, can now be seen to contain two components: An autonomous (not crisis-related) component, $\beta_{i,j}$, and a crisis-sensitive component, $\delta_{i,j} \times CRISIS_t$. The evolution of a firm's CAPM beta, therefore, is allowed to depend on developments in Asia in an easily estimable way.[9]

It also is possible to interpret model 2 not solely as a time-varying extension of the (one-factor) CAPM but as a multifactor asset-pricing model (Merton, 1973; Ross, 1976; Fama and French, 1992; Campbell and Cochrane, 1999).[10] This interpretation is valid only if the two "factors" represented by the U.S. market excess return and the product of the Asian and U.S. market excess returns essentially are uncorrelated. It turns out in our sample that they are: The pair-wise correlations of the two independent variables are -0.155 and -0.173 for the two crisis proxies that we investigate. Thus, we also can interpret our results as estimates of the sensitivity of firms' excess returns to first, U.S. excess market returns, and second, the product of Asian and U.S. excess market returns. If the two factors were completely uncorrelated, the estimated coefficient on the first (non-time varying component) would, in fact, be the (unconditional) CAPM beta.

We use weekly observations of returns from the beginning of 1997 through the end of 1998 to estimate the model. We examine two alternative crisis variables, each of which is the excess log return (in dollar terms) on an Asian portfolio,

[8] We chose one-year windows because the firm characteristics of interest in our second-stage regression are observed annually. At the same time, one-year windows allow for a sufficient number of weekly observations for the estimation of the parameters of interest in the asset-pricing model.

[9] An increase in the firm's exposure to market risk has two effects on the parameter $\delta_{i,j}$. First, investors demand higher future returns, which causes $\delta_{i,j}$ to increase. Second, an increase in $\delta_{i,j}$ implies that, for given expected future cash flows, the firm's stock price will decline initially. This reinforces the comovement of the firm's stock with the crisis variable.

[10] Campbell and Cochrane (1999, p. 7) demonstrate the equivalence of a one-factor conditional CAPM and an unconditional multifactor CAPM in which the two factors are the excess market return and the product of the excess market return and a time-varying predetermined financial variable (the log dividend-price ratio in their case). In our case, the financial variable is the value of an Asian-crisis proxy.

$R_{Asia} - R_f$, where the risk-free return R_f, is the weekly return on the three-month U.S. Treasury bill. The first crisis variable we use is the excess log return on the FT/S&P Actuaries Pacific Stock-Market Index excluding Japan. The second crisis variable is the excess log return on the J.P. Morgan Emerging Local Markets Index Plus (Asia), an index of money-market securities issued in Asian emerging markets.[11]

We assume $Cov[\varepsilon_{i,t},\varepsilon_{i,s}] = 0$ for $t \neq s$, where $t,s = 1,...,T$, that is, the error terms are serially uncorrelated. We allow the error terms to be heteroskedastic across equations (i.e., firms) but not over time. This means that $Var[\varepsilon_{i,t}] = \sigma_i^2$ need not be identical for all i ($i = 1,...,N$) but they are assumed to be identical for all t ($t = 1,...,T$).[12] We adopt the seemingly unrelated regression (SUR) approach but impose a diagonal covariance matrix for the error terms. That is, for all t ($t = 1,...,T$), we impose the restriction $Cov[\varepsilon_{i,t},\varepsilon_{j,t}] = 0$, $i \neq j$ ($i, j = 1,...,N$). This restriction is appropriate because regressors are identical across the set of firm-specific regression equations that are used in the regression model 2.[13]

We test for serial correlation using the Ljung-Box (1978) test statistic with a standard lag length of $floor\ (4\ (T/100)^{2/9})$, where *floor* means rounded down to the nearest integer. The null hypothesis of no serial correlation could not be rejected. We used the heteroskedasticity-consistent covariance matrix proposed by White (1980). Both the *t*-tests and the Wald tests are based on heteroskedasticity-consistent standard errors.[14] The White correction procedure deals with heteroskedastic residuals in the most general way. In estimating equation 2, the procedure controls for any form of heteroskedasticity that goes beyond the cross-sectional differences in the variances σ_i^2 ($i = 1,...,N$) for which the model controls directly.[15]

Results of First-Stage Regression

Although a larger sample size would have been desirable, we were able to obtain reliable geographical breakdowns of sales data for only 39 of the S&P 100 firms (see Tables 1 and 2 for the list of firms, their sales exposures to Asia, their financial leverage, and sample statistics for these variables). We obtained total-return series and calculated weekly excess returns (i.e., the return in excess of the return on a three-month T-bill) for each firm's stock and for the following (all in U.S. dollar terms): the S&P 500 index; the FT/S&P Actuaries World Stock-Market Index Pacific excluding Japan; and the J.P. Morgan Emerging Local Markets Index Plus (Asia), a money-market return index.[16] The cumulative total U.S. dollar returns for these three risky portfolios and for the riskless T-bill are plotted in Figure 2.

Estimates of the "autonomous" part of each firm's market sensitivity. The estimated coefficients, $\hat{\beta}_{i,1997}$ and $\hat{\beta}_{i,1998}$, would retain their interpretation from the traditional one-period CAPM framework as measures of the comovement of the excess return of firm i with the only relevant (priced) factor—the excess return of the market (the S&P 500) in year j—only under the null hypothesis that there is no crisis-related time variation in this factor loading (i.e., estimated sensitivity). If there is some systematic relationship between movements in our proxies for the Asian crisis and the way the stock-market prices our sample firms' equity returns, however, then our beta coefficients will quantify only the "autonomous component" of firms' exposures to market risk. Table 3, which appears on page 24, displays our estimated autonomous β-coefficients, where we have used the excess log returns on the FT/S&P Asian stock-market index in the first two columns and the excess log returns on the J.P. Morgan money-market index in the last two columns, respectively, as a proxy for the Asian crisis.

Estimated autonomous β-coefficients generally are significant statistically in Table 3, with 73 of 78 coefficients significant at the 10-percent level when using the Asian stock-market index. The Wald statistic for the overall fit of the 39-equation system of estimated β-coefficients is significant statistically in each year, as well. Significant Wald statistics for tests of cross-sectional and intertemporal differences indicate that there is a great deal of firm heterogeneity and intertemporal variation with respect

[11] For a list of the variables employed in this paper, see Appendix B.

[12] We control for the effects of a possible violation of the assumption of time-invariant variances by using White's (1980) heteroskedasticity-consistent covariance matrix.

[13] In this case, accounting for potential contemporaneous correlation across equations offers no efficiency gain. See Greene (1997, p. 676).

[14] See Greene (1997, p. 488) on how to implement covariance matrices for nonspherical disturbances in Wald tests.

[15] This implies that the White-corrected standard errors in the SUR framework are identical to those that would result if the equations in model 2 were White-corrected individually.

[16] See Appendix A for a detailed data description.

Table 1

Sample Firms
Panel A: Industry Classifications

Firm	Ticker Symbol	SIC Code	Primary Industry Classification
Aluminum Company of America	AA	33	Primary Metal Industries
American Express Company	AXP	61	Nondepository Financial Institutions
AMP Incorporated	AMP	36	Electronic & Electric Equipment
Avon Products, Inc.	AVP	28	Chemicals & Related Products
BankAmerica Corporation	BAC	60	Banks & Banking Services
Baxter International, Inc.	BAX	38	Instruments & Related Products
Boeing Company	BA	37	Transportation Equipment
Bristol-Myers Squibb Company	BMY	28	Chemicals & Related Products
Champion International Corporation	CHA	26	Paper & Related Products
Coca Cola Company	KO	20	Food & Related Products
Computer Sciences Corporation	CSC	73	Business Services
Delta Air Lines, Inc.	DAL	45	Air Transportation
Eastman Kodak Company	EK	38	Instruments & Related Products
Fluor Corporation	FLR	16	Heavy Construction
General Electric Company	GE	36	Electronic & Electric Equipment
General Motors Corporation	GM	37	Transportation Equipment
H.J. Heinz Company	HNZ	20	Food & Related Products
Halliburton Company	HAL	16	Heavy Construction
Homestake Mining Company	HM	10	Metal Mining & Related Services
Intel Corporation	INTC	36	Electronic & Electric Equipment
International Business Machines Corporation	IBM	35	Industrial Equipment & Machinery
International Paper Company	IP	26	Paper & Related Products
Johnson & Johnson	JNJ	28	Chemicals & Related Products
Mallinckrodt, Inc.	MKG	28	Chemicals & Related Products
McDonald's Corporation	MCD	58	Dining & Drinking Places
Merck & Co., Inc.	MRK	28	Chemicals & Related Products
Merrill Lynch & Co., Inc.	MER	62	Security & Commodity Brokers
Minnesota Mining & Manufacturing Company	MMM	32	Stone, Clay & Glass Products
Mobil Corporation	MOB	29	Petroleum & Coal Products
Monsanto Company	MTC	28	Chemicals & Related Products
National Semiconductor Corporation	NSM	36	Electronic & Electric Equipment
Northern Telecom Limited	NT	36	Electronic & Electric Equipment
Oracle Corporation	ORCL	73	Business Services
PepsiCo, Inc.	PEP	20	Food & Related Products
Pharmacia & Upjohn, Inc.	PNU	28	Chemicals & Related Products
Ralston-Purina Company	RAL	20	Food & Related Products
Schlumberger Limited	SLB	13	Oil & Gas Production
United Technologies Corporation	UTX	37	Transportation Equipment
Xerox Corporation	XRX	38	Instruments & Related Products

Table 1

Sample Firms
Panel B: Values of the Variables *SALES* and *DEBT*

Firm	*SALES* 1996	*SALES* 1997	*DEBT* 1996	*DEBT* 1997
Aluminum Company of America	0.172	0.167	0.126	0.111
American Express Company	0.083	0.078	0.060	0.066
AMP Incorporated	0.193	0.198	0.039	0.033
Avon Products, Inc.	0.156	0.154	0.047	0.045
BankAmerica Corporation	0.092	0.079	0.062	0.054
Baxter International, Inc.	0.164	0.145	0.223	0.303
Boeing Company	0.335	0.301	0.146	0.161
Bristol-Myers Squibb Company	0.104	0.098	0.066	0.085
Champion International Corporation	0.000	0.000	0.314	0.351
Coca Cola Company	0.218	0.226	0.069	0.047
Computer Sciences Corporation	0.055	0.062	0.156	0.176
Delta Air Lines, Inc.	0.028	0.024	0.147	0.116
Eastman Kodak Company	0.154	0.160	0.039	0.045
Fluor Corporation	0.095	0.108	0.001	0.064
General Electric Company	0.045	0.043	0.181	0.153
General Motors Corporation	0.021	0.009	0.176	0.186
H.J. Heinz Company	0.119	0.121	0.265	0.271
Halliburton Company	0.158	0.166	0.045	0.096
Homestake Mining Company	0.192	0.379	0.125	0.202
Intel Corporation	0.302	0.289	0.031	0.016
International Business Machines Corporation	0.194	0.194	0.122	0.168
International Paper Company	0.112	0.106	0.237	0.267
Johnson & Johnson	0.123	0.128	0.070	0.052
Mallinckrodt, Inc.	0.077	0.053	0.169	0.182
McDonald's Corporation	0.119	0.133	0.278	0.265
Merck & Co., Inc.	0.089	0.074	0.048	0.052
Merrill Lynch & Co., Inc.	0.062	0.062	0.123	0.147
Minnesota Mining & Manufacturing Company	0.181	0.175	0.064	0.077
Mobil Corporation	0.217	0.259	0.096	0.084
Monsanto Company	0.069	0.087	0.144	0.184
National Semiconductor Corporation	0.337	0.338	0.132	0.111
Northern Telecom Limited	0.074	0.070	0.153	0.125
Oracle Corporation	0.143	0.142	0.000*	0.065
PepsiCo, Inc.	0.116	0.066	0.344	0.246
Pharmacia & Upjohn, Inc.	0.146	0.137	0.051	0.038
Ralston-Purina Company	0.148	0.119	0.300	0.392
Schlumberger Limited	0.295	0.315	0.062	0.088
United Technologies Corporation	0.129	0.119	0.086	0.076
Xerox Corporation	0.006	0.006	0.314	0.317

* Zero value is due to rounding.

Table 2

Summary Statistics for *SALES* and *DEBT*

1996 1997	Minimum	Median	Mean	Maximum	Standard Deviation
SALES	0	0.123	0.136	0.337	0.083
	0	0.121	0.138	0.379	0.094
DEBT	0	0.123	0.131	0.344	0.093
	0.016	0.111	0.141	0.392	0.098

to market risk exposures of large firms. Results are similar when excess log returns on the J.P. Morgan money-market index are used as the crisis proxy. We also tested the significance of firm-specific intercepts in model 2. Wald test statistics of 37.60 (for the stock-market index) and 34.98 (for the money-market index) indicate that firm-specific intercepts are—as a group—not significant statistically, and thus, are excluded from the regression equation.

Estimates of δ and the "crisis-sensitive" components of firms' market sensitivity. The first-stage regression produces estimates of δ for each firm i during each year, $\hat{\delta}_{i,1997}$ and $\hat{\delta}_{i,1998}$ ($i = 1,...,39$). It is important to note that our proxy variables for the Asian crisis are *negative*, on average, during 1997 and 1998 for the FT/S&P stock-market index and during 1997 for the money-market index. Therefore, to translate the regression coefficients into a measure of the influence of the Asian crisis on the β-coefficients, we multiply $\hat{\delta}_{i,1997}$ and $\hat{\delta}_{i,1998}$ by the mean of the Asian crisis variable during the respective year. The products of the means and the respective regression coefficients are displayed in Table 4. They are estimates of the "crisis-sensitive" component of the firm's CAPM beta.

Table 4 indicates that 23 of the 78 estimates of a possible crisis-sensitive component of firm exposures to market risk were different from zero at the 10-percent confidence level when the Asian stock-market index was used as the crisis proxy. This is about three times as many as would be expected by pure chance. When the Asian money-market index was used as the crisis proxy, 20 of 78 estimates were significant at the 10-percent level. Thus, we conclude that it is possible to split large firms' exposures to market risk (their CAPM betas) into autonomous and crisis-sensitive components. As in the systems of estimated autonomous betas discussed above, the Wald statistics all are significant statistically in tests for the over-all fit of the crisis-sensitive components model, as well as for cross-sectional and intertemporal differences in crisis-sensitive components in firm returns.

By adding the crisis-sensitive component of a firm's beta from Table 4 to its corresponding autonomous part, presented in Table 3, we obtain the average total sensitivity of the firm to the S&P 500 return. This is the firm's beta measured at the mean of the respective crisis variable. Table 5, which appears on page 26, provides summary statistics that condense the information presented in Tables 3 and 4.

It is important to note that the sensitivity of the excess log return of firms we examine are unlikely to increase (or decrease) all at the same time. This is because a firm's beta measures the position of this firm relative to the index.[17] For example, if we had the entire set of S&P 500 firms in our sample, some betas necessarily would decrease if others increased. If greater sales exposure to Asia added to the firm's sensitivity to market risk, the betas of the firms that have a higher-than-average fraction of their sales to Asia would increase in response to the crisis while betas of those firms with below-average sales to Asia would have to decrease. The index average corresponds to the S&P 500's "average" firm.

Using the FT/S&P index returns in Tables 3 and 4, for example, we estimate the components of Minnesota Mining and Manufacturing's (3M) average total U.S. market (i.e., S&P 500) sensitivity during 1997 to be 0.497 and 0.111, for a total of

[17] A common influence of the Asian crisis on all firms' contributions to market risk would not lead to changes in the firms' betas. This is because a common impact (by definition) would affect all firms equally, leaving their positions relative to the market index unchanged.

Table 3

Estimates of the Autonomous Component of CAPM Betas, $\beta_{i,j}$

Firm	Asian Stock-Market Index 1997	Asian Stock-Market Index 1998	Asian Money-Market Index 1997	Asian Money-Market Index 1998
Aluminum Company of America	0.810***	0.640***	0.764***	0.659***
American Express Company	1.229***	1.778***	1.297***	1.785***
AMP Incorporated	1.152***	0.095	0.850***	0.124
Avon Products, Inc.	0.733**	1.397***	1.004***	1.384***
BankAmerica Corporation	0.817***	1.245***	1.057***	1.232***
Baxter International, Inc.	1.113***	0.541***	1.041***	0.523***
Boeing Company	0.307*	1.187***	0.252	1.215***
Bristol-Myers Squibb Company	1.181***	0.945***	1.241***	0.946***
Champion International Corporation	0.783***	1.336***	0.814***	1.333***
Coca Cola Company	1.381***	1.085***	1.377***	1.083***
Computer Sciences Corporation	1.097***	0.788**	0.667*	0.781**
Delta Air Lines, Inc.	1.033***	1.203***	0.925***	1.184***
Eastman Kodak Company	0.721**	0.077	0.555*	0.087
Fluor Corporation	0.949***	1.198***	1.070***	1.184***
General Electric Company	1.111***	1.351***	1.242***	1.334***
General Motors Corporation	0.827***	1.112***	0.813***	1.117***
H.J. Heinz Company	0.892***	1.818***	0.592*	1.858***
Halliburton Company	1.051***	0.508***	0.921***	0.505***
Homestake Mining Company	0.261	0.353	0.273	0.408
Intel Corporation	1.074***	0.642**	0.988***	0.637***
International Business Machines Corporation	0.897***	0.816***	0.729***	0.828***
International Paper Company	0.901***	0.753***	1.147***	0.764***
Johnson & Johnson	1.035***	0.485***	1.218***	0.470***
Mallinckrodt, Inc.	0.290*	0.807***	0.391**	0.810***
McDonald's Corporation	0.614***	0.941***	0.589***	0.953***
Merck & Co., Inc.	1.188***	0.858***	1.135***	0.866***
Merrill Lynch & Co., Inc.	1.928***	2.055***	1.990***	2.052***
Minnesota Mining & Manufacturing Company	0.497***	0.650***	0.578***	0.635***
Mobil Corporation	0.916***	0.680***	1.011***	0.712***
Monsanto Company	0.829***	0.789	0.727**	0.781
National Semiconductor Corporation	1.467***	1.039**	1.252***	1.077**
Northern Telecom Limited	1.450***	1.604***	1.348***	1.602***
Oracle Corporation	0.774*	1.844***	0.383	1.885***
PepsiCo, Inc.	0.785***	0.367*	0.698**	0.386**
Pharmacia & Upjohn, Inc.	0.981***	0.807***	0.924***	0.806***
Ralston-Purina Company	0.857***	0.787***	0.803***	0.799***
Schlumberger Limited	1.065***	1.563***	0.745***	1.588***
United Technologies Corporation	0.803***	1.160***	0.746***	1.161***
Xerox Corporation	0.973***	0.981***	0.728***	0.988***
Wald statistic	1122***	916.2***	770.1***	930.1***
Wald statistic (cross-sectional differences)	119.2***	139.5***	92.80***	136.7***
Wald statistic (intertemporal differences)	74.86***		66.49***	
Observations per firm	52	52	52	52

*/**/*** significant at 10/5/1 percent levels (*t*-tests are two-tailed).

Table 4

Estimates of the Crisis-Sensitive Component of CAPM Betas, $\delta_{i,j} \times CRISIS_t$, at Annual Means of *CRISIS*

Firm	Asian Stock-Market Index 1997	Asian Stock-Market Index 1998	Asian Money-Market Index 1997	Asian Money-Market Index 1998
Aluminum Company of America	0.043	0.012**	0.046	–0.031
American Express Company	–0.097	0.005	–0.085	–0.011
AMP Incorporated	0.100	0.001	0.211***	–0.032
Avon Products, Inc.	0.112	–0.007	–0.088	0.020
BankAmerica Corporation	0.057	0.020	–0.099*	–0.001
Baxter International, Inc.	0.074*	–0.012***	0.075*	0.029**
Boeing Company	0.171***	0.026**	0.114	–0.051*
Bristol-Myers Squibb Company	–0.104**	0.002	–0.084	–0.003
Champion International Corporation	0.150	0.008	0.059	–0.004
Coca Cola Company	–0.010	0.008	–0.003	–0.004
Computer Sciences Corporation	–0.045	–0.002	0.206**	0.009
Delta Air Lines, Inc.	–0.169**	–0.009*	–0.027	0.029**
Eastman Kodak Company	0.044	0.019*	0.110	–0.025
Fluor Corporation	0.027	–0.010*	–0.051	0.024
General Electric Company	0.017	–0.011**	–0.061*	0.028*
General Motors Corporation	–0.009	–0.007	0.003	0.000
H.J. Heinz Company	–0.062	0.004	0.128	–0.048
Halliburton Company	–0.074	0.000	0.032	0.003
Homestake Mining Company	0.077	0.019	0.032	–0.076*
Intel Corporation	0.108*	–0.007	0.100	0.010
International Business Machines Corporation	0.155**	–0.001	0.167**	–0.012
International Paper Company	0.124	0.006**	–0.068	–0.018*
Johnson & Johnson	0.052	–0.010*	–0.071	0.024*
Mallinckrodt, Inc.	–0.023	–0.001	–0.065*	–0.003
McDonald's Corporation	0.080*	0.005	0.053	–0.017
Merck & Co., Inc.	–0.026	0.002	0.016	–0.011
Merrill Lynch & Co., Inc.	–0.066**	0.019**	–0.066	–0.012
Minnesota Mining & Manufacturing Company	0.111***	–0.008*	0.013	0.023**
Mobil Corporation	–0.057***	0.013**	–0.079***	–0.046***
Monsanto Company	–0.060	0.025	0.024	–0.011
National Semiconductor Corporation	0.136	0.016	0.182	–0.055
Northern Telecom Limited	0.010	–0.011	0.059	0.011
Oracle Corporation	0.340	0.013	0.378	–0.055
PepsiCo, Inc.	0.043	0.010	0.068	–0.029*
Pharmacia & Upjohn, Inc.	0.002	0.000	0.031	0.001
Ralston-Purina Company	–0.027	0.012	0.015	–0.023
Schlumberger Limited	–0.023	0.006	0.158**	–0.032
United Technologies Corporation	0.070	0.002	0.065	–0.002
Xerox Corporation	0.026	0.010*	0.143**	–0.016
Wald statistic	91.21***	72.72***	77.04***	67.03***
Wald statistic (cross-sectional differences)	87.13***	71.95***	76.89***	65.37***
Wald statistic (intertemporal differences)	91.04***		64.42***	
Observations per firm	52	52	52	52

*/**/*** significant at 10/5/1 percent levels (*t*-tests are two-tailed); tests refer to the actual regression coefficients, not to the displayed products between the regression coefficients and the annual means of the crisis variable.

Table 5

Summary of Results from First-Stage Regressions

1997 1998	Asian-Crisis Index	Minimum	Median	Mean	Maximum	Standard Deviation
Autonomous component of CAPM beta, $\beta_{i,j}$	Stock market	1.78×10^{-1}	9.18×10^{-1}	9.09×10^{-1}	1.72	3.17×10^{-1}
		1.35×10^{-1}	8.86×10^{-1}	9.49×10^{-1}	1.97	4.45×10^{-1}
	Money market	2.13×10^{-1}	8.42×10^{-1}	8.82×10^{-1}	1.81	3.51×10^{-1}
		1.50×10^{-1}	8.87×10^{-1}	9.55×10^{-1}	1.98	4.45×10^{-1}
Crisis-sensitive component of CAPM beta, $\delta_{i,j} \times CRISIS_t$, (at annual means of crisis index)	Stock market	-1.49×10^{-1}	2.09×10^{-2}	1.86×10^{-2}	2.92×10^{-1}	8.25×10^{-2}
		-1.35×10^{-2}	2.47×10^{-3}	3.77×10^{-3}	2.60×10^{-2}	1.05×10^{-2}
	Money market	-1.15×10^{-1}	2.17×10^{-2}	2.20×10^{-2}	3.54×10^{-1}	9.90×10^{-2}
		-7.88×10^{-2}	-6.19×10^{-3}	-9.02×10^{-3}	3.24×10^{-2}	2.55×10^{-2}

0.608. The comparable estimates for 1998 are 0.650 and -0.008, for a total of 0.642. As shown in the tables, all of these component estimates for 3M were statistically significant in the first stage regression.

THE ECONOMETRIC MODEL: SECOND-STAGE REGRESSIONS

We attempt to identify the link between the Asian crisis and changes in firms' exposures to U.S. stock-market risk—their CAPM betas—in a second-stage regression. To this end, we gathered firm-specific data on sales to the Asian region as a proportion of total sales, on (the book values of) long-term debt and total assets, and on industry participation. In the first-stage regression approach discussed above, we used two different excess-return measures as proxies for the Asian crisis, one based on the FT/S&P stock-market index and the other derived from the J.P. Morgan money-market index. We run the second-stage regression using the results of each set of first-stage regressions as dependent variables.

Outline of Second-Stage Estimation Procedure

This section details the second step in our estimation procedure, which examines the channel by which the Asian crisis affected the sample firms' stock-market betas. We do this by regressing the coefficients $\hat{\delta}_{i,1997}$ and $\hat{\delta}_{i,1998}$ obtained from the asset-pricing model on a set of firm characteristics. These firm attributes are sales to the Asian region, weighted by the firm's leverage, and the firm's primary industry affiliation.

Sales exposure to Asia and the sensitivity of CAPM betas to the Asian crisis. The estimated crisis-sensitive coefficients from equation 2, $\hat{\delta}_{i,1997}$ and $\hat{\delta}_{i,1998}$, serve as the dependent variables in the second-stage regression. We have one estimate for each of the 39 firms in each year, for a total of 78 observations. A standard pooled-time-series cross-section approach can be applied because the stochastic nature of the dependent variable is controlled for by the error term of the regression equation.

The explanatory variable of interest in this regression is *SALES* × *DEBT*, the product of the firm's ratio of sales to Asia (*SALES*) and its ratio of book value of long-term debt to total assets (*DEBT*). The reason for weighting *SALES* with *DEBT* is that an equity holder's actual returns are influenced by the firm's capital structure, that is, the relative amounts of debt and equity the firm has. The key point is that leverage (i.e., debt) amplifies the risk of the firm's underlying cash flows as perceived by the equity holders. See the insert on the next page for more details.

To avoid simultaneity problems in the regression model, the observations for *SALES* and *DEBT* were lagged by one calendar year. There also are indicator variables in the regression corresponding to the two-digit SIC codes of the sample firms' main lines of business. No intercept is included and one of the 19 industry indicator variables is excluded (SIC 73). This is appropriate because we do not assume a common shift in the stock-market

LEVERAGE EFFECTS ON BETA

To illustrate the effect of leverage on an equityholder's returns more formally, let the exposure of a firm's underlying cash flows to variations in the cash flows of all firms in the market be denoted the firm's "assets beta," or β_{assets}, and let the firm's stock-market risk exposure, or "equity beta," be written as β_{equity}. Finally, let the market comovement of returns on the firm's debt be denoted its "debt beta," or β_{debt}. Then, if D is the book value of the firm's debt and E is the book value of equity, it is possible to express the firm's equity beta as follows (see Brealey and Myers, 1996, pp. 213-17):

$$\beta_{equity} = \frac{D+E}{E}\beta_{assets} - \frac{D}{E}\beta_{debt}$$

If the firm had no debt, its asset beta and its equity beta would be identical. Given positive amounts of debt in its capital structure, however, a firm's equity beta is larger than its asset beta.

betas due to the Asian crisis for the reason discussed above.[18]

The second-stage regression equation reads:

$$(3)\qquad \hat{\delta}_{i,j} = \theta \times SALES_{i,j-1} \times DEBT_{i,j-1} + \sum_{k=1}^{K} \phi_k I_{i,k} + \omega_{i,t}$$

with $j = 1997$ for $t = 1,...,52$ and $j = 1998$ for $t = 53,...,104$. The regressor $I_{i,k}$ is an industry indicator variable that takes on unit value if the corresponding firm belongs to industry k ($k = 1,...,K$; $K = 18$), and is equal to zero otherwise.

The *SALES* variable is measured with error because firms reported regional sales on a nonstandardized basis during our sample period.[19] Thus, we used an instrumental-variables (IV) approach, which entails ranks of the leverage-weighted sales numbers as instruments for the actual leverage-weighted sales numbers.[20] We tested for the presence of serial correlation using the Ljung-Box (1978) test statistic with unit lag length.[21] The *t*-values and the Wald test statistics are based on White (1980) corrected standard errors.[22] We also provide bootstrapped Student's *t* intervals and bootstrap-*t* intervals. While the first type of bootstrap intervals makes use of the Central Limit Theorem, the latter is distribution-free.[23]

Results of Second-Stage Regressions

Second-stage regression results are presented in Tables 6 and 7. Our most important finding is that firms' leverage-weighted sales exposure to Asia is significantly related to changes in their CAPM betas.

Sales exposure to Asia and the sensitivity of CAPM betas to the Asian crisis. Tables 6 and 7 provide the results from our second-stage regression. We regress the crisis-sensitive coefficients $\hat{\delta}_{i,1997}$ and $\hat{\delta}_{i,1998}$, ($i = 1,...,39$) on firm characteristics. The first two columns of Table 6 show results using the Asian stock-market index as the crisis proxy, while the results in the last two columns were generated using the money-market index. The table indicates that the impact of the Asian crisis on the crisis-sensitive component of the firm's CAPM beta was associated significantly with leverage-weighted sales exposure to the region. To interpret the negative coefficients, recall that the average excess returns on the stock-market index during both 1997 and 1998, as well as the average money-market excess return during 1997, were negative. The negative coefficients on the variable *SALES* × *DEBT*, when multiplied by the average excess return on the crisis proxy, indicate that a worsening of the crisis increased the crisis-sensitive component of a firm's beta, the greater the firm's leverage-weighted

[18] We tested for the statistical significance of the intercept and found it insignificant for both crisis variables. The *t*-values were equal to –1.243 (stock-market index) and –1.309 (money-market index).

[19] For details on the collection of the sales data, see Appendix A.

[20] On the use of ranks as instrumental variables in regressions, see Greene (1997, pp. 440-42).

[21] We use the Ljung-Box test because it is related to the heteroskedasticity and autocorrelation consistent covariance matrix proposed by Newey-West (1987). This correction would be particularly important if serial correlation were a problem, although we did not find any evidence of serial correlation for either Asian excess-return series.

[22] A White (1980) test is not well-defined for our regression because the number of regressors in the matrix needed for the test exceeds the number of observations. This also is true after eliminating linearly dependent regressors.

[23] For details on these two bootstrapping procedures, see Efron and Tibshirani (1993, pp. 158-62).

Table 6

Effects on Crisis-Sensitive Coefficients, $\delta_{i,j}$

Independent Variable	Asian Stock-Market Index Coefficient	Asian Stock-Market Index *t*-Statistic	Asian Money-Market Index Coefficient	Asian Money-Market Index *t*-Statistic
SALES × DEBT	-1.813×10^{2}	–2.620***	-2.563×10^{2}	–2.729***
SIC 10	1.153×10	3.268***	–3.442	–0.351
SIC 13	4.170	2.165**	-1.127×10	–1.406
SIC 16	-5.752×10^{-1}	–0.239	–4.323	–0.969
SIC 20	4.903	1.689*	5.941	1.590
SIC 26	–5.986	–1.499	–6.951	–1.486
SIC 28	2.996×10^{-1}	0.169	-4.842×10^{-1}	–0.142
SIC 29	-2.940×10^{-1}	–0.055	7.087	2.076**
SIC 32	2.172	0.448	7.452	1.527
SIC 33	-8.886×10^{-1}	–0.625	–2.249	–1.191
SIC 35	–2.534	–1.203	–4.263	–0.987
SIC 36	1.008	0.477	-2.855×10^{-1}	–0.081
SIC 37	5.544×10^{-1}	0.288	8.932	1.913*
SIC 38	7.761×10^{-1}	0.190	–3.833	–0.750
SIC 45	3.247	2.128**	-1.549×10	–1.190
SIC 58	7.534	3.076***	1.401×10	2.767***
SIC 60	-5.784×10^{-1}	–0.136	–1.728	–0.254
SIC 61	–3.488	–0.693	1.102×10	2.980***
SIC 62	–3.452	–1.403	–4.371	–2.095**
R^2	0.147		0.179	
Wald statistic (SIC codes)	325.8***		85.96***	
Number of observations	78		78	

*/**/*** significant at 10/5/1 percent levels (*t*-tests are two-tailed).

sales exposure to Asia. Similarly, we would expect a greater decline in the crisis-sensitive component of a firm's beta, the greater its leverage-weighted sales exposure, if the change in the crisis proxy were positive.[24] We control for industry effects in both regressions by including SIC code indicator variables representing each firm's primary industry classification.

Significance tests that are more reliable than the *t*-values displayed in Table 6 are shown in Table 7. For the regressor of interest, *SALES × DEBT*, we provide results from two alternative bootstrapping procedures that confirm our previous findings. A firm's leverage-weighted sales exposure significantly affected the impact of the Asian financial crisis on the firm's CAPM beta.

It is interesting to note that the two SIC codes identified by Pollard and Coughlin (1999, p. 39) as suffering the largest declines in real exports to East Asia during 1998—metallic ores and concentrates (SIC 10) and crude oil and natural gas (SIC 13)—were identified in our regression (using the stock-market index) as experiencing large and significant increases in the crisis-sensitive component of their market betas. For both crisis indexes, significant Wald statistics indicate that there were important industry effects even after controlling for the effect of leverage-weighted sales exposure to Asia on betas. We do not wish to emphasize these industry results, however, because the one-stage approach described in the next section does not find significant industry effects.

[24]The partial derivative of the expression $(1 + \lambda_{i,j} \times CRISIS_t)$ from equation 1 with respect to $\delta_{i,j}$ is $CRISIS_t/\beta_{i,j}$, where $\beta_{i,j}$ is the autonomous (i.e., crisis-insensitive) part of firm *i*'s CAPM beta in year *j*. As Table 3 shows, all $\beta_{i,j}$ ($i = 1,...,39$; $j = 1997, 1998$) were positive.

Table 7

Effects on Crisis-Sensitive Coefficients, $\delta_{i,j}$ (Bootstrapped Standard Errors)

Asian-Crisis Index	Regression Coefficient	Bootstrapped Student's *t* Interval	Bootstrap-*t* Interval
Stock market	-1.813×10^2	$+/-1.654 \times 10^2$	-1.382×10^2; 1.399×10^2
Money market	-2.563×10^2	$+/-2.133 \times 10^2$	-1.790×10^2; 1.818×10^2

95-percent confidence intervals, based on 2,500 draws.

A ONE-STAGE REGRESSION APPROACH

The two-stage regression procedure applied in this article can be aggregated into a one-stage procedure. Inserting equation 3 into equation 2 results in:

$$(4)\quad R_{i,t} - R_{f,t} = \beta_{i,j}\left(R_{m,t} - R_{f,t}\right) + \tilde{\theta} \times SALES_{i,j-1} \times DEBT_{i,j-1} \times CRISIS_t \times \left(R_{m,t} - R_{f,t}\right) + \sum_{k=1}^{K} (\phi_k I_{i,k}) \times CRISIS_t \times \left(R_{m,t} - R_{f,t}\right) + \tilde{\varepsilon}_{i,t}$$

with $j = 1997$ for $t = 1,...,52$ and $j = 1998$ for $t = 53,...,104$. We must drop one firm from the sample (Champion International) because it has zero values for the variable *SALES* (see Table 1, Panel B).

We estimated model 4 with a SUR approach that is similar to the estimation procedure we applied to model 2. In contrast to model 2 however, the regressors in model 4 are not identical across equations, which means that the efficiency of the regression can be improved by accounting for contemporaneous correlation across the equations in a SUR model.[25] Thus, we relaxed the previously imposed restriction of no contemporaneous correlation, that is, $Cov[\varepsilon_{i,t},\varepsilon_{j,t}] = 0, i \neq j$ $(i, j = 1,...,N)$, in favor of $Cov[\varepsilon_{i,t},\varepsilon_{j,t}] = \sigma_{i,j}$, $(i, j = 1,...,N)$. We then applied an iteration procedure in the estimation of the cross-equation covariance matrix.

The important results from this one-stage regression approach were similar to those of our preferred two-stage approach (not reported). Both approaches indicate that the leverage-adjusted sales exposure of our sample firms was significantly associated with changes in the firms' exposure to stock-market risk. Industry effects were not robust across the two approaches, however.

CONCLUSION

The Asian crisis affected consumers, investors, firms, and national economies around the world in many ways. Our analysis highlights one specific channel of influence on a sample of large U.S.-based firms. We find that the Asian crisis changed many of these firms' sensitivity to U.S. stock-market movements, that is, their CAPM betas. We find evidence that the link connecting the Asian crisis and changes in the stock-market risk exposure of our sample firms is the firms' leverage-weighted sales exposures to the crisis region.

As a firm's beta rises in response to a financial crisis, investors demand higher excess returns. This raises the firm's cost of equity capital. A firm could respond to this by reducing the share of its total sales that go to the crisis region or by decreasing its leverage. As a mitigating effect, some shrinkage in the ratio of Asian to total sales occurs automatically in the wake of such a crisis because the falling dollar value of sales constitutes a smaller fraction of total revenues. Some firms also may find it desirable to take actions to accelerate their shift away from the region or to reduce leverage.

Another mitigating effect in the stock market is the decrease in the weights of firms with increasing market betas. Even if the firm's expected earnings do not decrease, the firm's stock price must drop initially to generate the higher future returns investors demand due to the increase in the firm's

[25] See Greene (1997, pp. 674-76).

beta. A reduction in the share price reduces the firm's market capitalization, and therefore, its weight in the market index. This effect may be reinforced by a depression of the firm's expected earnings due to a decline in sales revenues from the crisis region.

REFERENCES

Black, Fischer, and Myron S. Scholes. "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy* (May-June 1973), pp. 637-54.

Brealey, Richard A., and Stewart C. Myers. *Principles of Corporate Finance*, 5th ed., McGraw-Hill, 1996.

Campbell, John Y., and John H. Cochrane. "Explaining the Poor Performance of Consumption-Based Asset-Pricing Models," NBER Working Paper 7237, July 1999.

Economic Report of the President (Washington, D.C.: United States Government Printing Office, 1999).

Efron, Bradley, and Robert J. Tibshirani. *An Introduction to the Bootstrap*, Chapman & Hall, 1993.

Fama, Eugene F., and Kenneth R. French. "The Cross-Section of Expected Stock Returns," *Journal of Finance* (June 1992), pp. 427-65.

Greene, William H. *Econometric Analysis*, 3rd ed., Prentice-Hall, 1997.

Jagannathan, Ravi, and Zhenyu Wang. "The Conditional CAPM and the Cross-Section of Expected Returns," *Journal of Finance* (March 1996), pp. 3-53.

Ljung, Greta M., and George E.P. Box. "On a Measure of Lack of Fit in Time Series Models," *Biometrica* (June 1978), pp. 297-303.

Merton, Robert C. "An Intertemporal Capital Asset Pricing Model," *Econometrica* (September 1973), pp. 867-87.

Newey, Whitney K., and Kenneth D. West. "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica* (May 1987), pp. 703-08.

Pollard, Patricia S., and Cletus C. Coughlin. "Going Down: The Asian Crisis and U.S. Exports," this *Review* (March/April 1999), pp. 33-45.

Ross, Stephen A. "The Arbitrage Theory of Capital Asset Pricing," *Journal of Economic Theory* (December 1976), pp. 341-60.

Schmid, Frank A. "Quality Spreads in the Bond Market," *Monetary Trends* (July 1999), < www.stls.frb.org/docs/publications/mt/1999/cover7.pdf > .

White, Halbert. "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity," *Econometrica* (May 1980), pp. 817-38.

Appendix A: DATASET

We analyze 39 of the 100 firms in the "Standard & Poor's 100" stock-market index for calendar years 1997 and 1998. We were limited to this subset of the S&P 100 because the other firms did not provide a sufficiently detailed breakdown of international sales. Our criterion for inclusion in the sample was that the firm's reports must allow us to calculate the ratio of sales to customers in Asia to worldwide sales (where both items are stated in U.S. dollar terms).

Sales data and balance-sheet information (long-term debt and total assets) were taken from each firm's annual reports. Data on consolidated sales (including all subsidiaries) for each firm were obtained from the income statement. Data on sales to customers in Asia were taken from the notes to the consolidated financial statements. For firms whose fiscal years did not coincide with the calendar year, the dollar amounts of sales and the end-of-fiscal year items were interpolated linearly to determine corresponding calendar-year values.

Construction of the SALES variable and lines of business

Data on international sales are provided in the notes to the consolidated financial statements in annual reports of U.S. corporations. Unfortunately, current reporting of international sales is not standardized. Many companies provide almost no detail beyond a breakdown into domestic and foreign sales. Only 39 of the S&P 100 firms provided enough country or area detail to allow us to calculate sales to the Asia/Pacific region. For example, nine companies included sales data for areas outside Asia as part of their totals for Asia/Pacific. We assumed that these firms combined sales totals from regions where they do relatively little business with regions where they are more active. We believe the non-Asian sales included in the Asia/Pacific totals are insubstantial.

A second problem is that firms that distinguish between sales to other departments or divisions within the organization (intracompany sales) and sales to unaffiliated customers (third-party sales) do not always do so on a consistent basis when presenting geographic breakdowns. In four cases, firms did not distinguish between intracompany and third-party sales on a geographic basis. For these four firms we know the total amount of intracompany sales on a consolidated (world wide) basis, but we do not know what percentage of sales to Asia were intracompany. In these four instances, the sum of sales totals to the various world regions will be greater than total sales (because intracompany sales have not been eliminated). Fortunately, intracompany sales averaged only 5.4 percent of third-party sales for these four firms.

The SIC two-digit code for each firm's main line of business was taken from the 10-K filings found on the Securities and Exchange Commission's EDGAR website < http://www.sec.gov/edgarhp.htm > ; 1998.

Financial Markets Data

Weekly total-return data for our 39 sample companies and the S&P 500 index for 1997 and 1998 were purchased from Standard & Poor's Compustat and DRI respectively. Weekly returns were calculated from Friday to Friday (closing prices), with adjustment for stock splits and dividends.

The risk-free rate was proxied by a strategy of investing in three-month Treasury bills each week, holding the bill one week, then rolling over into the new three-month T-bill. We collected the "on-the-run" three-month yield as well as the "off-the-run" yield of the bill issued the previous week for each week during 1997 and 1998 from the *Wall Street Journal*. Using these yields, we calculated prices. Three-month T-bill log returns were calculated as follows:

$$Log\left(\frac{P_{off,t}}{P_{on,t-1}} \quad r_{t-1}\right),$$

where $P_{off,t}$ is the price of the "off-the-run" issue at time t; $P_{on,t-1}$ is the price of the "on-the run" issue at time $t-1$; r_{t-1} is the value of the return index at time $t-1$ ($t - 1 > 0$); and r_0 equals 100.

We used the log returns on two indexes of Asian securities to measure the economic situation in Asia. These were the *FT/S&P Actuaries World Indices-Pacific Excluding Japan*, a Pacific-region stock-markets index, and the *J.P. Morgan Emerging Local Markets Index Plus* (*ELMI+*)−*Asia*, a Pacific-region emerging-markets money-market index. The FT/S&P World Indices are owned by FTSE,

International Limited, Goldman, Sachs & Co., and Standard & Poor's. These indices are compiled by FTSE International and Standard & Poor's in conjunction with the Faculty of Actuaries and the Institute of Actuaries. More information on the *J.P. Morgan Emerging Local Markets Index Plus* can be found in *J.P. Morgan's Emerging Markets Bond Index Monitor*. Both indexes are stated in U.S. dollar terms based on current exchange rates. Further details are provided below.

At year-end 1998, the *FT/S&P Actuaries World Indices-Pacific Excluding Japan* index included Australia (75 companies), Hong Kong (66), Indonesia (26), New Zealand (18), The Philippines (22), Singapore (41), and Thailand (35). Malaysia (106 companies as of September 25, 1998) was removed from the index on October 1, 1998, following the Malaysian government's introduction of investment exchange controls on September 1, 1998.

According to the ground rules for the construction and maintenance of the FT/S&P Actuaries World Indices, criteria for inclusion in the World Indices are the following:

- Direct equity investment by non-nationals must be permitted;
- Accurate and timely data must be available;
- No significant exchange controls should exist that would prevent the timely repatriation of capital or dividends;
- Significant international investor interest in the local equity market must have been demonstrated;
- Adequate liquidity must exist.

Companies whose business is that of holding equity stakes in other firms or other investments are not excluded necessarily. Equity-holding firms that are excluded include split-capital investment trusts and companies whose share price is a direct derivation of the values of underlying holdings, such as mutual funds. Only shares listed on a stock exchange are eligible for inclusion. Where a company does not list all its shares in an eligible class, or does not list an entire class, these partially listed or unlisted shares are not eligible. All securities comprising the bottom 5 percent of a country's market capitalization are excluded from the indices. A security is totally excluded if foreign investors are barred from ownership. Calculation of the U.S. dollar version of this index is explained in detail in the FT/S&P Actuaries World Indices Ground Rules at < http://www.ftse.com > .

The second financial-market index we used was the *J.P. Morgan Emerging Local Markets Index Plus* (*ELMI+*) (*U.S. Dollar Index*)-*Asia*. The ELMI+ tracks total returns for local-currency-denominated money-market instruments in 24 emerging markets. It is predominantly non-Latin America weighted and includes four regional composites, Asia (48 percent target weight as of January 29, 1999), Europe (18.26), Latin America (23.07), and Middle East/Africa (10.67). The ELMI+ employs a liquidity-sensitive weighting system, which uses exports plus imports as a base. Its portfolio consists of FX forwards, wherever possible, to represent a country's money markets.

A country is selected if it has been identified as an emerging market with an economy large enough to support significant capital flows, and has accessible liquid local-currency-denominated money-market instruments, either on- or offshore.

The Asian regional sub-index was used for this analysis. The countries included in the Asian regional composite are China (4.167 percent of the total index), Hong Kong (20.833), India (4.167), Indonesia (16.67), The Philippines (4.167), Singapore (20.833), South Korea (4.167), Taiwan (4.167), and Thailand (20.833). A special August month-end rebalancing was performed to account for Russia's removal from the index, and a special September 8, 1998, rebalancing was performed to account for Malaysia's removal.

Target weights are derived by applying a series of caps to the three-year, rolling, trade-weighted allocation for each country. Specifically, for countries with convertible currencies, the weight per country is limited to no more than 10 percent of the total index. For countries with nonconvertible currencies or impediments to investing onshore, the weight per country is limited to no more than 2 percent of the index. For each country subindex, a ladder of three instruments was constructed, by initially investing in one-, two-, and three-month instruments. Each month, the proceeds of the maturing instrument is reinvested in a new three-month instrument.

Appendix B: DEFINITION OF VARIABLES

Dependent Variables

The dependent variable in the asset-pricing regressions is each sample firm's weekly excess return, defined as the difference between the log total return on the firm's stock and the risk-free return. A second-stage regression uses parameter estimates from the first-stage regression as the dependent variable. This parameter is the crisis-sensitive component of each firm's estimated market beta, $\hat{\delta}_{i,j}$ ($i = 1,...,39; j = 1997,1998$).

Independent Variables

We used four total-return indexes with each return denoted as follows:

R_i: Log change in total-return index of firm i (weekly);

R_m: Log change in S&P 500 total-return index (weekly);

R_f: Log change in three-month T-bill total-return index (weekly);

R_{Asia}: Alternatively defined as log change in the FT/S&P total-return index or in the J.P.M. total-return index (weekly).

Firm-specific variables include the following:

$SALES_i$: Ratio of dollar amount of sales to customers in Asia to dollar amount of consolidated sales of firm i (annual);

$DEBT_i$: Ratio of book value of long-term debt of firm i to its total assets (end of calendar year; annual);

$SICx_i$: Indicator variable: 1 if the firm's main line of business falls into the two-digit SIC division x, 0 otherwise (annual).

Daniel L. Thornton is a vice president and economist at the Federal Reserve Bank of St. Louis. The author thanks John Duffy, Kevin Hoover, David Laidler, Alvin Marty, and Bennett McCallum for helpful comments. Jonathon Ahlbrecht and Stephen Majesky provided research assistance.

Money in a Theory of Exchange

Daniel L. Thornton

> *"In primitive traffic the economic man is awaking but very gradually to an understanding of the economic advantages to be gained by exploitation of existing opportunities of exchange...Consider how seldom it is the case, that a commodity owned by somebody is of less value in use than another commodity owned by someone else! And for the latter just the opposite relation is the case. But how much more seldom does it happen that these two bodies meet!...Even in the relatively simple and so often recurring case, where an economic unit, A, requires a commodity possessed by B, and B requires one possessed by C, while C wants one that is owned by A—even here, under a rule of mere barter, the exchange of the goods in question would as a rule be of necessity left undone."*
> — *Carl Menger, "On the Origin of Money,"* The Economic Journal *(June 1892), p. 242.*

> *"Money, I consider, is a device which facilitates the working of markets." Sir John Hicks,* A Market Theory of Money *(1989), p. 2.*

A major problem in monetary economics has been to introduce money into the economy in a way that: (1) explains how money arises endogenously, (2) explains why money is preferred to other methods of exchange, and (3) identifies the welfare gains associated with money's use. Money has been introduced by including it as an argument in consumers' utility functions or producers' production functions, assuming the existence of a welfare-reducing cash-in-advance constraint, assuming that it is a vehicle for making intergeneration transfers with no role in exchange, or simply assuming that money exists—although it is given no specific role to play.[1]

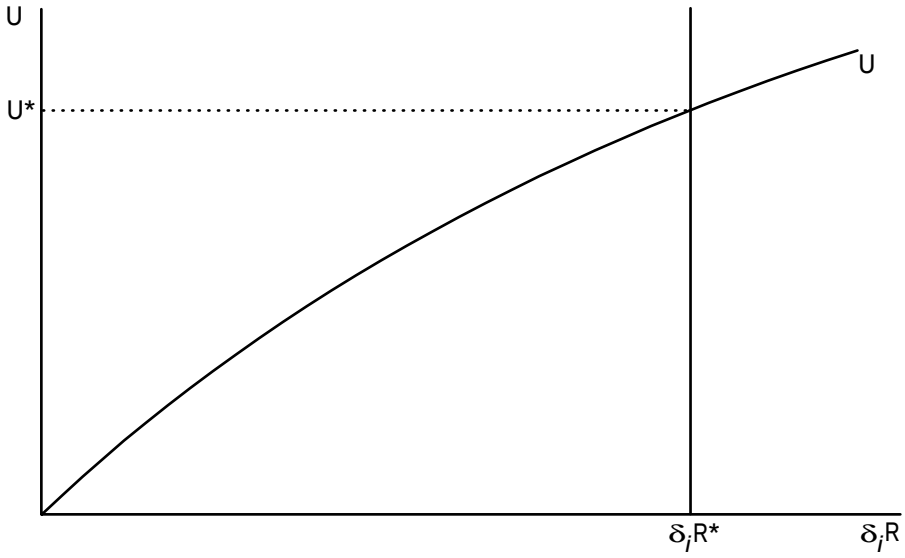
This paper develops a framework for assessing money's role and the welfare gains associated with its use. This framework shows how money reduces the resources necessary for exchange, thereby increasing both consumption and leisure. The welfare gains from trade are enhanced because the use of money promotes further trade and greater specialization. For expository purposes the analysis is linear; however, it is more correct to think of trade, money, and specialization as essentially evolving simultaneously, with the development of each reinforcing the development of the others. Nevertheless, I argue if there were no trade, there would be no need for money. To understand the role of money in an exchange economy it is not necessary to know the precipitous event that set off this evolutionary chain.

By showing how the use of money economizes on scarce resources, expands trade and promotes specialization, the analysis makes explicit Laidler's claim (1990, p. 47) that "...there is something of the nature of a public good about money" so that "...we should be very wary of treating the sum of its private products as its social product." In so doing, I show that to determine the welfare gains from money it is necessary to compare a monetary economy with economies that use alternative methods of exchange. The analysis has micro foundations, but contrary to the trend in macroeconomics, no formal, explicit general-equilibrium model of money is developed. Indeed, the analysis suggests that the obstacles to the formulation of such a model of money are considerable.

The analysis is presented in four sections. The first section develops a model of an autarkic economy and derives a measure of social welfare as a function of economic resources. In the second section, individuals are permitted to trade and several fundamental consequences of costly trade are derived and discussed. The third section argues that there only are three possible methods of effecting exchange: barter (simple and sequential), credit (simple and sequential), and money. Because economic agents have an incentive to choose the least costly method of effecting exchange, an analysis of the relative cost of each of these methods reveals why the world has been dominated by monetary, rather than barter or credit economies. The analyses in the first three sections provide the spring-

[1] See Hoover (1995) for a discussion of these and other approaches.

Figure 1



board for the fourth, a discussion of several interesting and difficult questions in monetary theory.

AN AUTARKIC ECONOMY

Many monetary analyses begin by modeling an economy with money and discuss money's implications. Since the purpose of this article is to better understand why money exists and how its use enhances welfare, it seems natural to start with an economy where there is no money. Moreover, since I argue that money's existence depends on trade, the natural starting place would seem to be a model of an autarkic economy. Consider an economy with N individuals and Q commodities. Each individual is endowed with a quantity of a non-depletable resource $\delta_i R^*$, where δ_i is the i^{th} individual's proportionate share, $0 \leq \delta_i \leq 1$, of the total economy-wide resource, R, which is fixed and given at R^*. Individuals are self-sufficient and maximize utility, where the i^{th} individual's utility function is

$$U_i(C_1^i, C_2^i, \ldots, C_Q^i, l^i). \tag{1}$$

C_j^i, $j = 1, 2, \ldots, Q$, denotes the quantities of the Q commodities consumed by the i^{th} individual and l denotes the amount of time devoted to leisure. Each individual produces these commodities via the following production functions,

$$C_j^i = f_j^i\left(\left[\delta_i R^*\right]_j, L_j^i\right) \quad j = 1, 2, \ldots, Q, \quad \text{and } i = 1, 2, \ldots, N, \tag{2}$$

where $[\delta_i R^*]$ denotes the physical quantity of the resource devoted to the production of the j^{th} commodity by the i^{th} individual and L_j^i, denotes the amount of the i^{th} individual's time devoted to the production of the j^{th} commodity. The i^{th} individual maximizes 1 subject to 2 and to the constraints

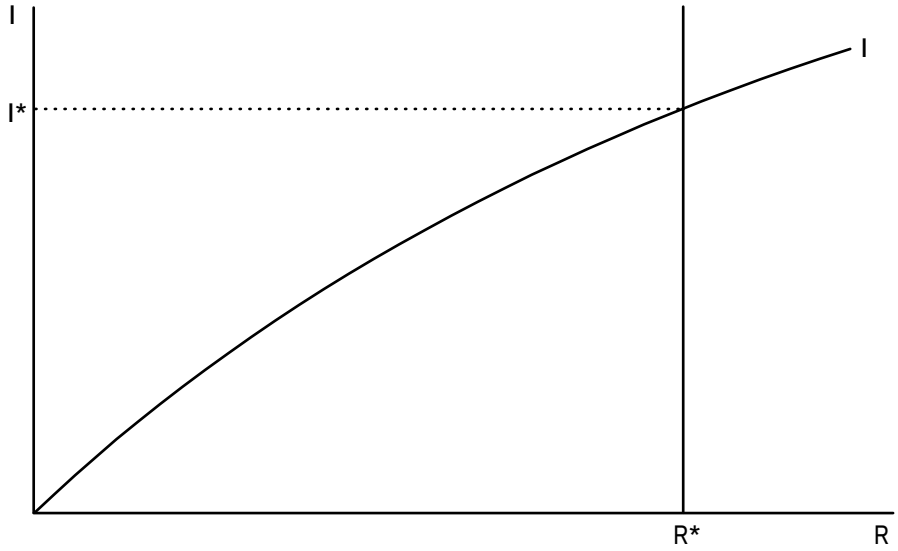
$$\sum_{j=1}^{Q} (\delta_i R^*)_j = \delta_i R^* \tag{3}$$

and

$$\sum_{j=1}^{Q} L_j^i + l^i = \Gamma. \tag{4}$$

to maximize utility, each individual must allocate resources, $\delta_i R^*$, among the production of the Q consumption goods and the total available time, Γ, among the productions of consumption goods and leisure. The solution to this optimization problem

Figure 2



results in the first-order conditions,

$$\frac{\partial f_j^i / \partial (\delta_i R^*)_j}{\partial f_k^i / \partial \left(\delta_i R^*\right)_k} = \frac{U_k^i}{U_j^i},$$
$$j,\ k,\ = 1, \ldots, Q,\ j \neq k$$

and

$$\left(\partial f_j^i / \partial L_j^i\right) = \frac{U_1^i}{U_j^i}, \quad j = 1, \ldots, Q.$$

These conditions are familiar. The first set requires individuals to allocate resources, $\delta_i R^*$, between the production of the goods that they consume by equating the ratio of the marginal utilities with the marginal rate of technical substitution for each pair of commodities consumed. The second set requires individuals to allocate time between the production of the goods that they consume and leisure by equating the marginal product of each good with respect to the labor devoted to its production to the ratio of the marginal utility of leisure to the marginal utility of that good.

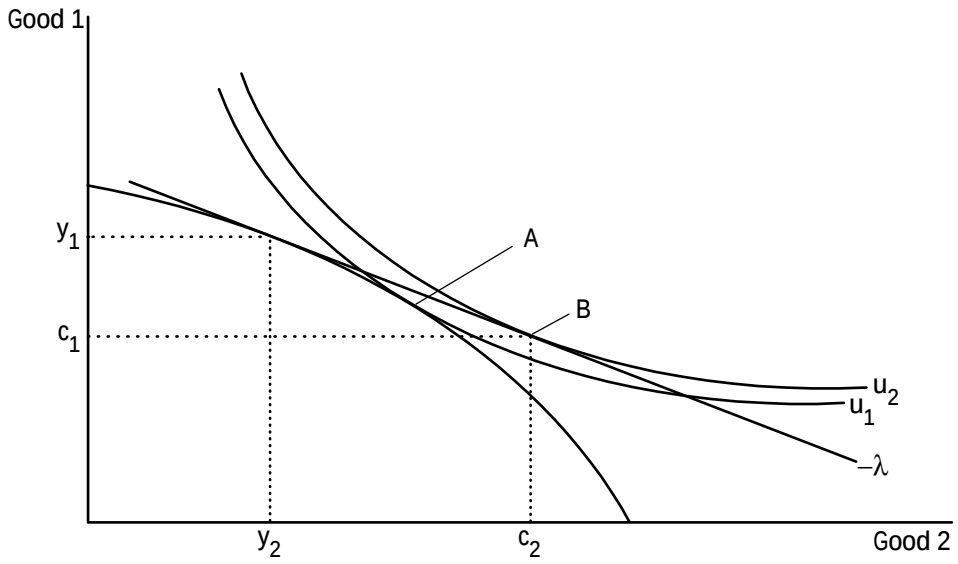
Let U_j^i, (C_1^*, C_2^*, ..., C_Q^*, l^*) denote the solution to the optimization problem for the i^{th} individual. Assume that utility is measured ordinally, i.e., each individual assigns a real number, n, to a particular level of utility such that $U' > U$, if $n' > n$. Under the usual assumptions about preferences and production, the i^{th} individual's maximum utility can be put into a monotonic relationship with that individual's resources, as illustrated in Figure 1. The point of interest is where U intersects the vertical line at $\delta_i R^*$. This is the point where the i^{th} individual maximizes his well-being given the state of technology, the available resources and time.

The simple aggregation of the ordinal utility measures over the N individuals yields an index of maximum utility for society as a whole, I, shown in Figure 2. Since this index is a linear combination of monotonically increasing functions in R, for a given distribution of resources, i.e., a given set of δ's, it is a monotonically increasing function in R as well. Society's well being is maximized given the state of technology, the available resources and time at the point where I intersects society's resource constraint, R^*. This is the point of maximum social welfare for the autarkic economy.

AN EXCHANGE ECONOMY

The purpose of this section is to illustrate the effects of trade and to show how transactions costs

Figure 3



reduce the benefits from trade, thereby limiting the extent of trade. The intent is not to develop a complete model of costly trade, explain the degree of specialization that one observes, or to specify the precise benefits from trade. For simplicity, leisure is omitted as an argument in the utility function and exchange and production require only time, not additional resources, R.

Figure 3 illustrates the autarkic optimum and the gains from trade. The point A is the autarkic, no-trade optimum for individual i. At point A the slope of the production frontier $-f_2'/f_1'$ is equal to the slope of the indifference curve, $-U_1/U_2$. The point marked B is the trade optimum, given the exchange ratio, λ. By producing more of good 1 and less of good 2, relative to autarky, the individual is able to reach a level of consumption that was infeasible without trade, thereby, increasing utility from u_1 to u_2. With trade the individual produces y_1 units of good 1 and y_2 of good 2 and consumes C_1 and C_2 units of good 1 and good 2, respectively.

Now consider the effects of costly trade. The analysis is kept simple by assuming that individual 1 wishes to maximize the utility function $U(C_1,C_2)$ and can produce these goods with the following technologies:

$$y_1 = f_1(L_1)$$

and

$$y_2 = f_2(L_2),$$

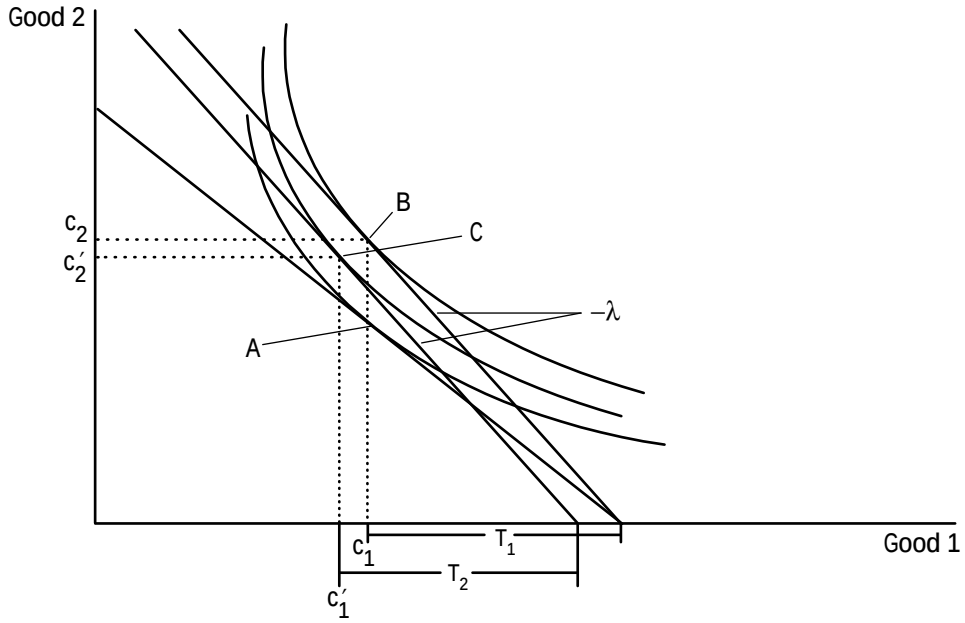
where y_1 and y_2 are the quantities of the two goods produced and L_1 and L_2 are the amounts of labor time devoted to the production of each of the two goods. Assume that the production technologies are linear, so that the marginal rate of technical substitution, f_2'/f_1', is constant. Further assume that individual 1 specialized in the production of good 1, which can be traded for good 2. Note that if exchange is costless, trade will be advantageous for any exchange ratio $\lambda > f_2'/f_1'$.

Assume that the transaction cost, i.e., the amount of time needed to trade, is fixed at Ω and does not vary with the volume of trade. Given this assumption and the others, the Lagrangian, L, can be written as

$$L = U(C_1, C_2) + \mu(C_2 - \lambda(f(\Gamma - \Omega) - C_1)). \quad (5)$$

Differentiating and solving the usual first-order

Figure 4



conditions, yields

$$\lambda \geq \frac{f_2'}{f_1'} + (f_1(\Gamma) - C_1^A)^{-1}\Omega,$$

(6) $$\frac{U_1}{U_2} = \lambda.$$

Equation 6 is the familiar condition that the marginal rate of substitution equals the exchange ratio.

The effect of costly trade on optimal consumption is illustrated in Figure 4. The autarkic, no-trade optimum is denoted by *A*. The costless trade optimum, for a given exchange ratio, λ, is denoted by *B*. The fixed-cost trade optimum for the same exchange ratio is denoted by *C*. Costly trade reduces the welfare gains from trade, as the consumption of both goods is smaller when trade is costly. Note that the volume of trade, T_2, is smaller when trade is costly than when trade is costless, T_1. The reason is that trade draws resources, in this specific example time, away from production. Indeed, if the cost of trade, Ω, is large enough, no trade will take place—the autarkic optimum will dominate the trade optimum.

The most important thing to notice, however, is that the exchange ratio that is required to achieve the costless trade outcome is larger when trade is costly. This is seen by noting that to a first-order approximation,

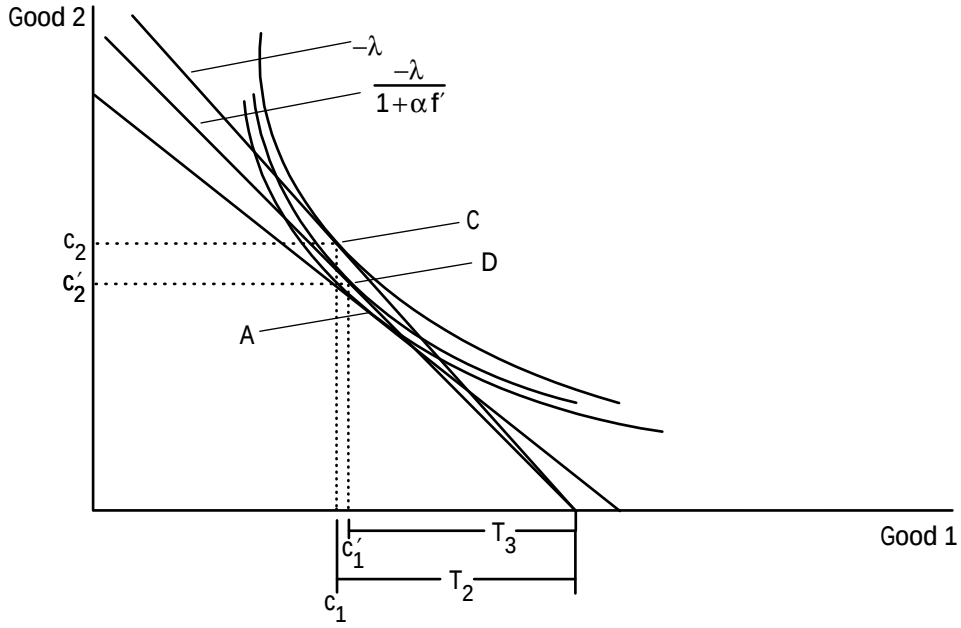
where C_1^A, is the optimum consumption of good 1 in the autarkic economy (see the appendix for details). When trade is costly ($\Omega > 0$), the individual must be compensated for the cost per unit of y_1 that must be given up to make the trade.

This effect of costly trade is more apparent when exchange costs vary with the volume of trade. The exact outcome depends on the assumption made about the nature of these trading costs, however, the basic effect of costly trade will be invariant to their nature. Hence, for simplicity, assume that the exchange cost, L_e, the time that is required to trade, is proportionate to the volume of trade, i.e.,

$$L_e = \alpha(f(L_1) - C_1), \quad \alpha > 0.$$ [2]

[2] Note that given this specification, the marginal exchange costs of using one more unit of time for exchange rather than production, i.e., $dL_e/dL_1 = \alpha f'$, depends on the marginal product of labor. The greater the marginal product of labor, the larger is the marginal cost of exchange. This fact may help explain why some individuals specialize in production and others specialize in marketing or exchange. For example, the stereotypical western *storekeeper* is someone who cannot hunt, rope, ride, or steal.

Figure 5



Again, the individual is assumed to specialize in the production of good 1. Moreover, the total amount of time, Γ, is used either in the production of good 1, L_1, or exchange, L_e, i.e., $\Gamma = L_1 + L_e$. The LaGrangian for this maximization problem is

$$L = U\left(C_1, \lambda(f(L_1) - C_1)\right) + \mu(\Gamma - L_1 - \alpha(f(L_1) - C_1)). \tag{7}$$

The first-order conditions are:

$$U_1 - \lambda U_2 + \mu\alpha = 0$$

$$\lambda U_2 f' - \mu - \alpha\mu f' = 0$$

$$\Gamma - L_1 - \alpha(C_1 - f(L_1)) = 0$$

Solving the first-order condition yields,

$$\frac{U_1}{U_2} = \frac{\lambda}{(1 + \alpha f')}. \tag{8}$$

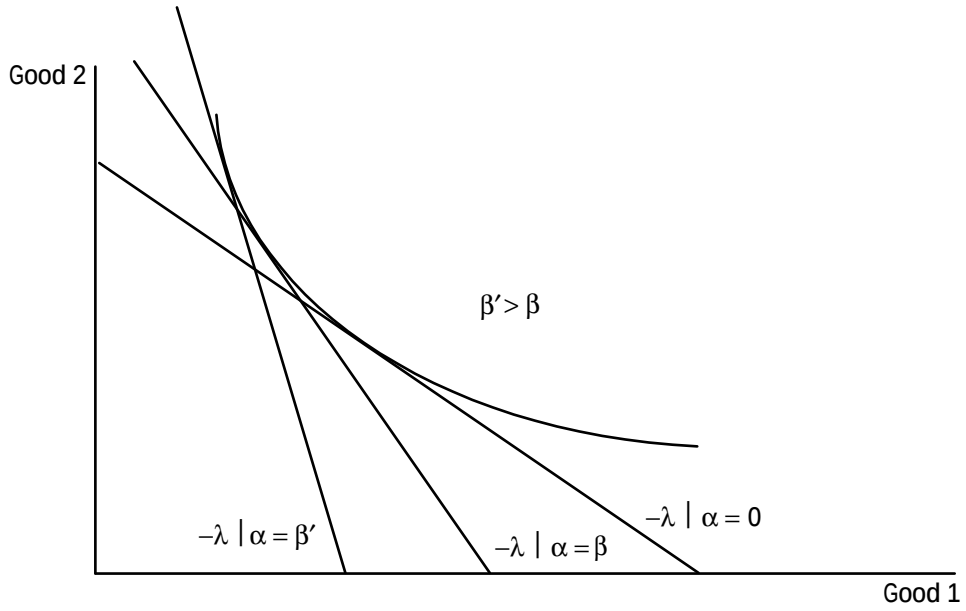
If $\alpha = 0$, this condition reduces to the previous one. When there are no exchange costs, an individual who specializes in the production of good 1 is better off trading whenever the exchange ratio, λ, is greater than or equal to the individual's fixed marginal rate of technical substitution, i.e., $\lambda > f_2'/f_1'$. When exchange costs vary with the volume of trade, however, the marginal condition for trade becomes $\lambda > (1+\alpha f_1')(f_2'/f_1')$.[3]

The effect of marginal exchange costs is illustrated in Figure 5. The point marked C is the same as that in Figure 4; namely, it is the optimal point assuming that the exchange costs are fixed. D denotes the optimum when exchange costs vary with the volume of trade. The effect of variable trade costs is to reduce the effective exchange rate for a given exchange ratio, λ. The gains from trade are smaller than when exchange costs are fixed and there is a corresponding reduction in the volume of trade.[4] All other things being the same, the volume of trade falls from T_2 to T_3. The trade optimum, for a given λ, is pushed closer to the autarkic no-trade optimum, which again is denoted by A. If the exchange costs are sufficiently high,

[3] The strict inequality is due to the fact that an individual must be compensated for the total amount of y_1 that must be given up to make the trade.

[4] Figure 5 is drawn on the assumption that the total cost of trade in this example is exactly equal to the fixed costs of trade in the previous one, i.e., $\alpha(f(L_1)-C_1) = \Omega$.

Figure 6



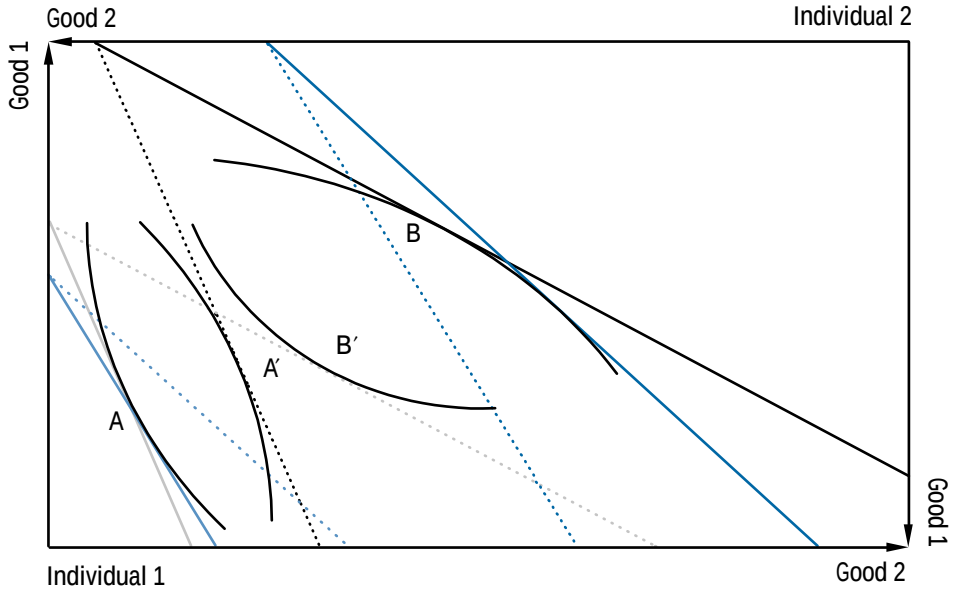
the autarkic optimum will dominate the trade optimum for a given λ.

Costly trade not only reduces the gains from trade, but more importantly, it increases the minimum exchange ratio that is required for the individual to benefit from trade. Define the *reservation exchange ratio*, *rer*, to be the minimum exchange ratio required for an individual to acquire the same level of utility that would be acquired under autarky. If exchange is costless, *rer* is simply an individual's marginal rate of technical substitution. Figure 6 shows *rer* for three assumptions about exchange costs: There is no cost of exchange, i.e., $\alpha = 0$, and two cases where exchange costs are positive, $\alpha = \beta > 0$ and $\alpha = \beta'$, $\beta' > \beta$. Figure 6 illustrates that the higher the exchange cost, the larger is *rer*, i.e., the more of good 2 that an individual who specializes in the production of good 1 must get to compensate for the cost of trade. Trade is advantageous only when the terms of trade are sufficiently favorable, i.e., λ is sufficiently large to compensate for the cost of trade.

All of the above conclusions were predicated on the assumption that the individual specializes in the production of good 1. Hence, it is important to see how costly trade affects the potential exchange between individuals where the comparative advantage is well defined. Assume that individuals 1 and 2 are able to produce both goods 1 and 2. Again, the production technologies are assumed to be linear and the solid gray and black lines, respectively, in Figure 7 denote their production frontiers. Individual 1 has a comparative advantage in the production of good 1; individual 2 has a comparative advantage in the production of good 2. The exchange ratio at which trade can take place, λ, must be between the slopes of the solid gray and black lines which, in the case of costless trade, represent individual 1's and 2's *rer*, respectively. The points *A* and *B* are optimal if each individual is self-sufficient. Point *A′* denotes a trading possibility where all of the gains from trade accrue to individual 2, while *B′* denotes a trade possibility where all of the benefits from trade accrue to individual 1.

Now assume that both individuals have marginal exchange costs, i.e., $\alpha_1 f_1'$ and $\alpha_2 f_2'$, which are positive but not necessarily equal. The effect of costly exchange is to raise the *rers* for both individuals. The solid light blue and dark blue lines represent the *rers* for individuals 1 and 2, respectively, when trade is costly. The dashed dark blue line is parallel to the solid light blue line and, hence, denotes the maximum benefits from trade when trade is costly if all the benefits from trade accrue to individual 2. Likewise, the dashed light blue line is parallel to the

Figure 7



solid dark blue line and, hence, denotes the maximum benefits from trade when trade is costly if all the benefits accrue to individual 1. The maximum benefits from trade are clearly less when trade is costly and diminish as the cost of trade increase.

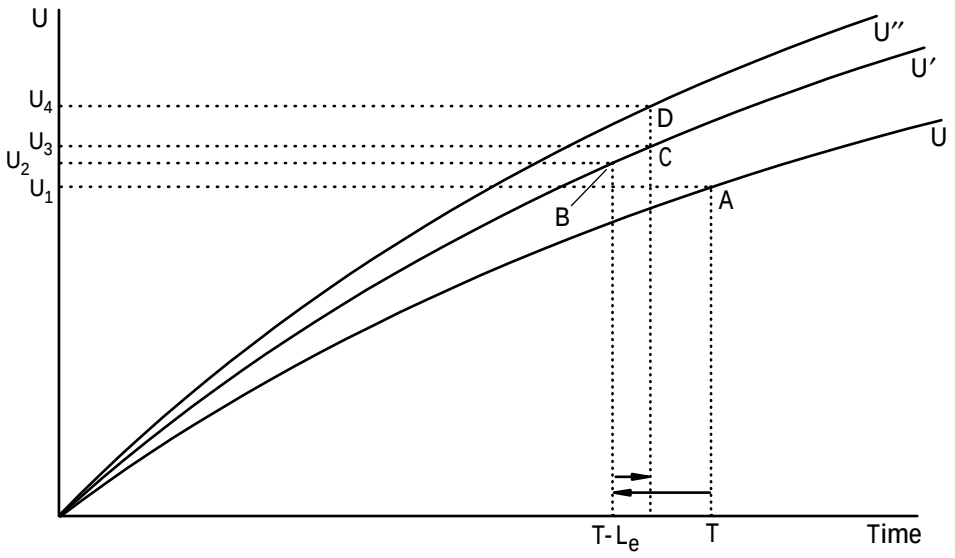
The important thing to note is that while the precise gains from trade for the two individuals depend on the respective size of the transactions costs, costly trade reduces the range of exchange ratios where trade is mutually advantageous and, therefore, the volume of trade. Moreover, the larger the exchange costs, the smaller the region where trade is mutually beneficial. Indeed, if the slopes of the solid light blue and dark blue lines were sufficiently large for either individual, no exchange ratio would exist where trade would be mutually advantageous—no trade would take place.[5] Costly trade reduces the feasible set of opportunities where trade is mutually advantageous.

The exchange ratio at which individuals trade and how much each benefits from trade depends on the relative costs of trade for both individuals, which in turn depends on strategic considerations that go well beyond the scope of this inquiry. For example, the above analysis assumed that the costs were borne by both traders and that there were no social arrangements for sharing the costs. Moreover, there is nothing in this analysis that ensures these individuals trade or that more trade takes place if the exchange costs are reduced. The conclusion that trade increases is inferred from noting that individuals have an incentive to engage in mutually advantageous trade up to the point where the marginal resource cost-of-trade equals the utility gains from trade. Anything that reduces exchange costs gives rise to this potential by expanding the feasible set of mutually beneficial trades.

The existence of exchange costs and the desire to reduce them has implications for the development of markets and, more generally, for their structure. Anything that reduces transactions cost encourages greater trade and specialization. At the same time, the benefits from specialization and trade encourage the use of the most efficient method of exchange. Of course, the catalyst for all of this is the heterogeneity that makes trade mutually advantageous. Reducing the cost of trade enhances welfare by (a) reducing the amount of resources that

[5] Since trading costs are positively related to the real volume of the goods traded, the *rers* will vary with the level of trade. This makes determining the exact amount of trade in costly trade environments very difficult. Moreover, it has been assumed that the trade costs are proportional to the volume of trade, but this need not be the case. Trading costs also may vary across individuals or goods. Regardless of how exchange costs are treated, the same fundamental conclusion emerges: The larger the exchange costs the less trade will take place.

Figure 8



must be devoted to exchange, freeing up resources for production (and/or leisure) and by (b) increasing the amount of trade that takes place, i.e., increasing the extent of the market.

A host of mechanisms have evolved to reduce the cost of exchange: bazaars, trading posts, retail establishments, brokers, agents, dealers, and other specialists and, most especially, money. Some of these have given way to more efficient methods of exchange. Others have not—at least not yet.

The effect of innovations, such as money, which reduce the cost of exchange, is illustrated in Figure 8. Point *A* denotes the level of utility for an individual who is self-sufficient. All time is spent in production or leisure. Point *B* denotes the level of utility associated with costly trade. The utility level is higher than for autarky despite the fact that some time, L_e, is used for exchange. Money reduces exchange costs, so that fewer resources are devoted to exchange, and there is more time for production (or leisure). This gain can be seen by comparing points *B* and *C*. In addition, the use of money increases welfare by expanding the set of feasible transactions and, thereby, increasing the volume of trade. This is illustrated by the difference between points *C* and *D*. The total gains from reducing exchange cost are illustrated by the difference in utility levels associated with points *B* and *D*. Applying the same aggregation analysis as before yields the implication that any reduction in exchange cost is welfare enhancing for society as a whole. This analysis makes it clear that by reducing the transaction cost, money expands the set of exchange ratios where trade is mutually advantageous. To this extent, the use of money expands the feasible set of transactions.[6]

The Implications of Costly Trade

The objective of the above analysis was to illustrate a role of money and the benefits from money's use, not to construct a general theory of trade. Indeed, the analysis says nothing about who trades with whom or how much trade takes place. Rather, it illustrates that trade is welfare enhancing and that costly trade reduces welfare (relative to costless trade), both by drawing resources from production or leisure and by limiting the extent of the markets. Money increases economic welfare by mitigating some of these costs. Moreover, welfare is enhanced even though money does not appear in the utility or production functions or is a prerequisite for trade through a cash-in-advance

[6] For the view that money does not increase the set of feasible transactions, see Ostroy, (1973), pp. 608-9.

constraint; nor are there legal restrictions requiring the use of money. Furthermore, there is nothing to rule out the possibility that some transactions are achieved using barter or credit. The analysis confirms Brunner and Meltzer's argument (1971, p. 804) that "...the private and social productivity of money are a direct consequence of the saving in resources that the use of money permits and of the extension of the market system that occurs because of the reduction in the cost of making exchanges."[7]

The conclusion that money enhances economic welfare by reducing exchange costs is independent of the market structure, so long as there is some degree of decentralization.[8] Of course, the exact nature and extent of the benefits from innovations that reduce exchange costs depend on such factors. Consequently, any attempt to quantify the benefits associated with innovations that reduce exchange costs is necessarily stylized: Specific results will depend on the assumptions made about the structure of markets, production technologies, the nature and extent of the exchange costs, and who bears them. Since my purpose is to gain insight into how money ameliorates exchange costs, it is essential to deal with these costs very generally. Before turning attention to the issue of exchange cost, however, several other implications of the above analysis for money are noted.

First, innovations that reduce the exchange cost of one individual can benefit others. This implication is clear from Figure 7. Instead of both individuals experiencing a reduction in exchange cost, assume that only one individual does. The effect still is to increase the feasible region of exchange ratios where exchange could take place. Consequently, an innovation that reduces the exchange cost of one individual can raise the utility of both.

Second, individuals have an incentive to use the least-costly method of exchange. The fact that some form of money has evolved in every society suggests that money is efficient relative to other methods of effecting trade. Moreover, that similar assets have functioned as money in very different societies suggests that certain assets seem to have a distinct advantage in reducing exchange costs.

Third, exchange costs limit the extent of trade and, hence, specialization and the use of money reduces these costs. Consequently, it is not surprising that the division and specialization of labor appear to have evolved simultaneously with the use of money.

Finally, and perhaps most importantly, the welfare gains from money can be obtained only by comparing a monetary economy with an economy that uses an alternative method of exchange. Moreover, the welfare gains from money will change as markets develop and economies become increasingly specialized. Consequently, the more highly specialized the economy and the greater the extent of trade, the larger the likely effects of disruptions to the supply of money will be. As we will see, this point has implications for the welfare costs of inflation.

It has been recognized for some time that economies with a medium of exchange are better off than if no such medium of exchange exists. Indeed, McCallum (1983a, p. 24) uses "the *traditional presumption* that an economy with a medium of exchange is more productive than it would be if no medium of exchange existed" to show that overlapping generations models of money (e.g., Wallace, 1983) have no role for money as a medium of exchange. The above analysis refines McCallum's point by explicitly showing how money necessarily enhances welfare by facilitating trade. Models that do not explicitly recognize this role of money are unlikely to capture money's essential feature.[9]

---

[7] Despite the large amount of resources devoted to market activity (bringing buyers and sellers together) economic analyses have focused on production and consumption. For an exception, see Hirshleifer (1973).

[8] Ostroy (1973) was one of the first to observe that the Walrasian general equilibrium market had no role for money because no trade takes place until the equilibrium set of *accounting prices* Patinkin (1965), is determined. This is why Meltzer (1995), Hicks (1989), and others argue that such models may be of limited use in understanding the role of money in the economy.

[9] Even models that explicitly capture the medium-of-exchange function of money do not necessarily capture the welfare enhancing properties of money noted here. For example, the shopping-time model of McCallum and Goodfriend (1987) or the money-in-exchange models of Dornbusch and Frenkel (1973), Benhabib and Bull (1983), and Fischer (1986) explicitly recognize the medium-of-exchange function. In the latter models, however, the metric for measuring aggregate welfare provides no motive for exchange. In these models welfare is measured by aggregate consumption, which is given by,

$$C = f(k)\left[1 - v(m)\right]; \quad 0 \leq v < 1,\ v' < 0,$$

where $v(m)$ is the proportion of total output, $f(k)$, that is used in exchange. Since C for $v > 0$ is always less than C for $v = 0$, exchange does not enhance economic welfare. Hence, there is no motive for exchange and, consequently, no motive for money as a medium of exchange. The problem is analogous to that of Tobin (1965) and Mundell (1971), where per-capita output is maximized when money holdings are zero.

This is not true of the search-theoretic models of money (e.g., Jones, 1976; Oh, 1989; Kiyotaki and Wright, 1989, 1991, 1993; Trejos and Wright, 1993; Johri, 1994) that explicitly model money's role as a medium of exchange. In these models, money facilitates exchange by ameliorating the search costs associated with the *double coincidence of wants* essential for barter. Implications of some of these models have been born out experimentally (e.g., Duffy, 1998).

Before discussing why money exists and why money dominates barter and credit as a medium of exchange, one final point should be made. Specifically, the welfare gains from money discussed above are those associated with the real stock of money, i.e., money's purchasing power. If money is held solely to facilitate transactions, a larger real money stock means that more transactions are facilitated and, hence, the welfare gains are larger relative to the next best method of exchange; therefore, society's welfare should increase with the equilibrium stock of real money. Changes in the nominal stock of money, however, do not necessarily result in an increase in the equilibrium stock of real money. Given classical neutrality and Archibald and Lipsey's (1958) invariance principle, *ceteris paribus*, increases in the stock of nominal money may have no significant effect on economic welfare.[10]

EXCHANGE COSTS WITH ALTERNATIVE METHODS OF EXCHANGE

The previous section showed why anything that reduces the exchange cost is welfare enhancing. Such innovations enhance welfare by reducing the quantity of society's scarce resources devoted to exchange, freeing up resources, time for production, or time for leisure, and by enabling society to achieve a greater extent of specialization and trade. I inferred that money is one such innovation without carefully defining what money is. This section takes up this issue. Specifically, money is defined as a commodity that is a generally acceptable medium of exchange.

The essential feature of an exchange economy is that individuals trade the commodity that they have, commodity *j*, for one that they want, commodity *k*. The essential point is that there are only three possible methods of exchange: barter, money, or credit. In an exchange economy, trade must take place with one of these methods.[11] Which of these methods is used depends on their relative costs in effecting exchange. Hence, the analysis of money necessarily requires an analysis of the relative costs of barter, credit, and money in exchange.

In discussing the relative costs of these alternative methods of exchange, it is important to distinguish between simple barter (trading commodity *j* for commodity *k*) and multistage or sequential barter (trading commodity *j* for commodity *h* and trading commodity *h* for commodity *k*). Money always entails a sequential transaction (trading commodity *j* for *m* and trading *m* for commodity *k*). Hence, a monetary transaction can be thought of as a sequential barter transaction that involves a particular commodity, *m*. When *m* becomes generally acceptable, it is money.[12]

It also is important to distinguish between simple and sequential credit transactions. A simple credit transaction involves trading commodity *j* for a promise of some commodity (*k*, *j*, *m* or some other commodity) in the future. Hence, a simple credit transaction is just an intertemporal barter transaction.

A sequential credit transaction involves trading commodity *j* for an IOU and trading the IOU for commodity *k* or perhaps another IOU.[13] According to this definition, a sequential credit transaction is a particular form of a sequential barter transaction where the intermediate commodity is an IOU. If a particular IOU were *generally acceptable*, it would be *money*. Simple credit transactions have been used to effect exchange for a long time, e.g., trade credit. IOUs have even circulated as a form of local currency for relatively short periods of time. There are a number of reasons, however, why money initially took the form of tangible commodities and not IOUs. Indeed, I will argue later that the use of credit for the purpose of facilitating trade is due to the existence of money. Hence, while credit can supplement money in effecting exchange, it will not supplant it. In this section, however, I only will consider the question of how well credit can facilitate exchange. For this purpose, sequential credit is required.

It is important to distinguish between costs that are independent of the method of exchange and those

[10] See Bullard (1999) for a survey of the evidence for monetary *neutrality* and *superneutrality*.

[11] McCallum (1985) has also made this point.

[12] In the search literature on money, *general acceptability* is achieved in various ways. For example, in Oh's (1989) model where individuals search randomly, a dominant medium of exchange emerges due to the assumption that one commodity has the largest subjective probability of trade. He shows that if traders try to minimize the number of encounters that result in the desired trade, e.g., commodity *j* for commodity *k*, this commodity will emerge as the dominant medium of exchange. Money and barter coexist because barter occurs when an individual who has commodity *j* and wants commodity *k* just happens to meet an individual with commodity *k* and who wants commodity *j*.

[13] Brunner and Meltzer (1971) consider what they termed a barter-credit economy. In their discussions, credit is simply bartered for goods. That is, one person gives the other an IOU for the goods that the former wants. But this implicitly assumes that the latter person wants the IOU and not some other commodity that is desired for current consumption. The case where credit is used in barter transactions is discussed later.

that vary with the method of exchange. That is, it is essential to differentiate between costs that money can ameliorate and those that it cannot.[14] To this end, exchange costs, i.e., all costs associated with making the exchanges of two or more commodities between two or more individuals, are categorized as either information costs or noninformation costs.

Noninformation costs are the packaging, handling, and other assorted costs associated with getting a commodity from the point of final production to the point of final consumption. Such costs are independent of whether the transaction is achieved with barter, credit, or money. Hence, while such costs are essential for determining the extent of and the benefits from trade, they are immaterial for the broader question of why money is used and for determining the welfare gains from its use.[15]

Information costs vary with the method of exchange. Information costs are divided into *assurance* costs and *shopping* costs. Comparisons of the assurance costs associated with money, barter, and credit explain why money dominates sequential barter or sequential credit in exchange. Comparisons of the shopping costs associated with money, barter, and credit add to this explanation.

Assurance Costs

Consider first the case of multistage transactions. This is when individuals trade a commodity they have for one that they currently do not. This process will continue until they obtain the desired commodity. Such multistage transactions require that individuals obtain assurance that they will be able to obtain the desired commodity, *k*. Broadly speaking there are two distinct, although not mutually exclusive, categories of costs associated with obtaining this assurance. The first of these I term *verification costs*. Verification costs, which are discussed extensively by Brunner and Meltzer (1971) and Alchian (1977), are the costs of verifying the characteristics and attributes of the good received. These costs include the costs of inspection, measuring, perfecting property rights, etc.[16]

The second category of costs I call *value-determination costs*. Value-determination costs are the costs associated with determining the value or worth of the commodity received. The value of the commodity is the number of units of it that must be traded for *x* units of the desired commodity, *k*. Clearly, a poor-quality commodity is less valuable than a high-quality commodity. Nevertheless, knowing the quality of the product does not necessarily mean that you know its exchange value, hence, it is useful to treat these information costs as separate and distinct.

A sequence of barter transactions that ultimately leads to the acquisition of commodity *k*, requires verification costs at each stage in the sequence.[17] Hence, the verification costs associated with sequential barter could be considerable, especially if a large number of intermediate transactions are required. Money economizes on verification costs relative to sequential barter for two reasons. First, money has relatively low verification costs. Indeed, other things being the same, the commodity with the lowest verification cost will emerge as money (Jevons,1875).[18] Second, the use of money means at most two transactions, *j* for *m* and *m* for *k*, are required.

The verification costs of sequential credit transactions are likely to be high. If one person exchanges commodity *j* for an IOU of Mr. Smith, he has no difficulty in verifying that this is the IOU of Mr. Smith; after all, he watched Mr. Smith write it.[19] When he attempts to trade Mr. Smith's IOU for commodity *k*, however, the verification costs for the next person are likely to be significant. There may be considerable difficulty in verifying that this IOU is the promise of a particular Mr. Smith.[20] Moreover, even if

[14] In this context, it is somewhat arbitrary to assume where production ends and exchange begins; however, Hirshleifer (1973) suggests treating transportation costs as part of production costs.

[15] Of course, there could be second-order effects. Specifically, an innovation to the method of exchange could significantly increase the extent of the markets, which may in turn reduce transportation costs if there were economies of scale in transportation. Such innovations also could foster innovations in the transportation industry.

[16] Alchian (1977, p. 134) argues that it is the low verification costs alone that make a commodity money.

[17] See Jones (1976), Oh (1989), Kiyotaki and Wright (1993) and Trejos and Wright (1993).

[18] Jevons' observation is explicitly modeled by Jones (1976) and Oh (1989), who assume that one good is more *in demand* than other goods. Specifically, they take the "subjective expected transaction costs"—the time spent searching for complimentary trading partners—to be the number of encounters one anticipates before completing a single trade.

[19] Search-theoretic models abstract from the problem of credit by assuming that chance meetings of individuals have a Poisson distribution, so that the probability of the same individuals meeting twice is infinitesimally small relative to the probability of meeting once.

[20] Of course, methods have been developed to deal with such verification costs, but the costs still may be rather high relative to the verification costs of money.

one knew it was the IOU of a particular Mr. Smith, one might not know Mr. Smith's willingness and ability to honor the obligation. Even if the IOU is collateralized, all subsequent holders of the IOU must evaluate and perfect their interest in the collateral. Because of the costs associated with such activities, it seems likely that the verification costs of sequential barter in IOUs will be higher than the verification costs of sequential barter in commodities. Consequently, if money has lower verification costs than sequential barter, it must have lower verification costs than sequential credit as well.

The individual also must determine the value of the commodity, i.e., the number of units of the commodity k (or h) that they can get for x-units of the commodity j. The value-determination costs of sequential barter are high because it requires that the individual know up to $Q(Q-1)/2$ relative prices. The problem associated with a multitude of prices is exacerbated if credit is used to effect trade. Credit instruments can be denominated in any one of the Q commodities for any one of the N individuals. Hence, the credit price of a commodity can vary across goods and individuals. Credit prices also can differ in other dimensions, such as the maturity of the contract, whether there is collateral and the nature and extent of the collateral. Given problems associated with asymmetric information, it seems that the cost of determining the value of IOUs is so high that it eliminates the possibility that a credit instrument—denominated in a nonmoney asset—could serve as an effective, generally acceptable medium of exchange, i.e., it could serve as money.[21] Money has significantly lower value-determination costs than either sequential barter or sequential credit because traders are required to know at most $Q-1$ money prices.

Value-determination costs also depend on the variability of the value of money. All other things being the same, money serves best as a medium of exchange when its value remains relatively stable. Because of the difficulty in determining the relative value of commodities, it would be surprising to find that a commodity whose value fluctuates considerably—relative to other commodities—serves as a generally accepted medium of exchange, i.e., serves as money.

Maintaining the stability of money's value over long periods of time is important for what Jevons (1875, pp. 5-6) called the *standard of value* function of money. Because of money's role as a medium of exchange, and hence a source of generalized purchasing power, it is convenient to denominate credit contracts in terms of money. That is, the existence of money facilitates the use of credit. Indeed, as Hoover (1988) has noted, and I will argue later in more detail, money appears to be essential for credit. Variation in the value of money can have a significant, detrimental effect on money's standard of value function.[22] I will argue later that this effect may be larger than the effect of variation in the value of money on money's function as a medium of exchange.

Shopping Costs

If an individual who has commodity j and wants commodity k runs into an individual who has commodity k and wants commodity j, is the result necessarily barter? Some insights into the answer to this question come from considering shopping costs. Shopping costs are of particular interest because these are the costs that money is ordinarily thought to ameliorate relative to simple barter. Shopping costs encompass a wide variety of costs, including costs associated with locating a seller or buyer, haggling over price, budgeting, and planning expenditures. Shopping costs that are of particular interest are those associated with the lack of a *double coincidence of wants* and those associated with not having a common *unit of account*. While the medium of exchange and unit of account functions are separable, the fact that most often the same good has performed both functions suggests that this arrangement is efficient.

Trade is a planned activity. Sellers seek buyers and buyers seek sellers. Trade that occurs as the result of chance meetings is rare.[23] Indeed, special-

[21] Note that if credit were to supplant money, credit contracts would have to be denominated in something other than money. Moreover, it should be clear from this discussion that credit is less likely to arise as a general method of trading goods in primitive economies, where the costs of acquiring information are relatively high. Credit is more likely to exist as the primary method of exchange in advanced societies where information costs are relatively low.

[22] Jevons (1875, p. 6 and p. 12).

[23] In the search-theoretic literature, e.g., Jones (1976), Oh (1989), Kiyotaki and Wright (1993), and Trejos and Wright (1993), trade results from chance encounters. In such settings, money arises because it minimizes the costs of transactions due to a double coincidence of wants. Traders bump into each other randomly and engage in simple barter if there is a double coincidence of wants, they trade goods for money if one of the traders has money, or they do nothing. Recently, search models of money have allowed for the development of longer-term relationships, see Corbae and Ritter (1998).

ists who are particularly efficient at verifying and determining the value of particular products arise. In some cases, these specialists make markets for specific commodities by dealing in them. Matching *buy* and *sell* orders and managing inventories is difficult if a whole host of commodities is traded for the commodity(s) that the specialist deals in. Hence, the specialist's function is more efficient if transactions are carried out in a single commodity. Efficiency is further enhanced if *bid* (offers to buy) and *ask* (offers to sell) prices are quoted in this same commodity.

The efficiency of the market is increased if all participants agree to use the same medium of exchange and if this commodity also serves as the unit of account.[24] The elimination of the double coincidence of wants reduces the time for buyers and sellers to locate each other. Money also reduces the time spent haggling over price if everyone agrees to quote prices in the same commodity and if that commodity is accepted generally in exchange. For analogous reasons, planning and budgeting are made easier if a single commodity is used as both the medium of exchange and the unit of account.

Let us now reconsider the intriguing question that began this section: *If an individual who has commodity j and wants commodity k runs into an individual who has commodity k and wants commodity j, is the result necessarily barter?* In the search-theoretic literature of money (e.g., Jones, 1976; Oh, 1989; Kiyotaki and Wright, 1993; and Trejos and Wright, 1993), the answer is unequivocal. Yes! In these models, money and barter coexist because chance encounters sometimes result in a double coincidence of wants. The scarcity of barter is related inversely to the probability of such encounters. In economies where monetary exchange is well established, however, barter is scarce because seeking a double coincidence of wants becomes increasingly inefficient with the increased use of money and specialization (Jevons, 1875, p. 3). Individuals who possibly could barter might bump into each other and never know it. Furthermore, they would not care. In highly specialized monetary economies, barter is motivated more by tax considerations or thin markets for peculiar goods (such as second-hand goods—goods become more heterogeneous as they get used—e.g., trading a used car in on the purchase of a new car) rather than by chance encounters.

Why Is Money Held?

The above analysis explains why money is the most efficient method of effecting transactions, it does not explain why money is held. The classical explanation of why money is held deals with the lack of synchronization of receipts and expenditures. Brunner and Meltzer (1971, p. 785, fn. 4) challenged this view, arguing

> It is easy to see why a "lack of synchronization" does not imply that money is used and held. Consider an economy that has neither a medium of exchange nor money. If there are no costs of acquiring information, differences in the timing of receipts and payments are adjusted by issuing verbal promises in exchange for goods and, later, delivering goods. More generally, in a barter-credit economy, commitments or promises to pay bridge the gap between receipts and payments.

Brunner and Meltzer are correct that the lack of synchronization does not explain why money is held and they also are correct in suggesting that costly information explains money's dominance in exchange. Money is held for only one reason—by its very nature the process of exchange takes time so that anything that functions as a medium of exchange must be held. This is so obvious that Brunner and Meltzer (1971, p. 804) refer to it as "trivial." Trivial though it is, this fact alone is sufficient to explain why money is held. While economists have dealt with time in a number of ways, it is convention to refer to things that are measured as a point in time as *stocks* and things that are measured through time as *flows*. Out of necessity, money is a stock. If the stock of money facilitates the flow of trade, it must exist before and after the trade takes place. In a barter economy, the goods that are traded in the interval from t to $t + 1$ must exist at time t. If, hypothetically, trade and consumption are permitted to occur simultaneously, consumption goods bartered at time t do not exist at $t + 1$. In the case of money (or sequential barter or credit),

[24] Niehans (1978) has argued that the medium-of-exchange and unit-of-account functions are wedded because money cannot ameliorate the pricing problem noted above unless money also is the unit of account. The wedding of the medium-of-exchange and unit-of-account functions enhances market efficiency also has been noted by White (1984, p. 711).

however, this is not the case. The nominal quantity of money held at $t + 1$ must be the same as that held at t. Hence, although money is continuously changing hands, it always is being held by someone—it is never consumed. This is not solely a characteristic of money; it is true of any asset.[25] Assets traded at time t exist at both t and at $t + 1$.

What then distinguishes money from any other asset? *Ipso facto* every asset is a store of wealth. To say money is a store of wealth is tautological! Being a store of wealth is not a defining characteristic of money or any other asset.[26] Money can be distinguished only from other assets by applying another criterion.[27] The important criterion for separating money from other assets is that money is an asset (or group of assets) that is generally acceptable as the means of trading goods—other assets are not. This characteristic distinguishes money from other assets and has a long tradition in classical monetary economics (e.g., Menger, 1892; Jevons, 1875; Brunner and Meltzer, 1971; and McCallum, 1983a, b, 1985). At different times and under different circumstances various assets have served as money. Nevertheless, some assets appear to have characteristics that have made them the predominant forms of money.

Money is distinguished from other assets by the function it performs; this is illustrated by a simple story from Jevons (1875):

> When Mr. Wallace was traveling in the Malay Archipelago, he seems to have suffered rather from the scarcity than the superabundance of provisions. In his most interesting account of his travels, he tells us that in some of the islands, where there was no proper currency, he could not procure supplies for dinner without special bargain, and much chaffering upon each occasion. If the vendor of fish or other coveted eatables did not meet with the sort of exchange desired, he would pass on, and Mr. Wallace and his party had to go without their dinner. It therefore became very desirable to keep on hand a supply of articles, such as knives, pieces of cloth, arrack, or sago cakes, to multiply the chance that one or other article would suit the itinerant merchant. (pp. 2-3).

The Wallace party undoubtedly kept a cache of articles that were most highly demanded by the natives. These articles are clearly stores of wealth, but the party's members would normally not have chosen to store their wealth in this form. Rather, these particular articles were held because they facilitated trade, i.e., because they reduced transaction costs. The Wallace party used these articles as a form of local currency. Whether an asset can be used as money depends solely on whether it generally is *held* to facilitate exchange.

Finally, it is worth emphasizing that money is unique among assets not solely because it facilitates the exchange of goods for consumption. Money also facilitates the exchange of other nonmoney assets as well. Individuals typically do not exchange shares of stock for acres of land even when the person who has land wants stock and vice versa. Rather, land is traded for money and the money for stock. More importantly, it is generally convenient to denominate credit contracts in units of money rather than bushels of wheat, acres of land, or other commodities or assets. Generally speaking, assets are held for the myriad of reasons that individuals accumulate wealth. In contrast, money is primarily held because of its low cost in effecting transactions. Money is distinguished from other assets in that it is the only asset that is a generally held medium of exchange. Because of this, it also is the standard of value.

ISSUES IN MONETARY THEORY

The above analysis has implications for several interesting issues in monetary theory such as the origins of fiat money, the asset demand for money, the relationship between money and credit, the buffer-stock notion of money demand, the welfare benefits of money, and the welfare costs of inflation. I will now discuss each of these issues in turn.

Fiat Money

Until now, money implicitly has been a tangible real commodity, i.e., a commodity money, or a claim to such. In modern monetary economies, however, money is typically paper currency with no intrinsic value. A question that has troubled monetary econ-

[25] Assuming, of course, that assets do not depreciate.

[26] Hicks (1989, p.42) also has made this point.

[27] There are several criteria for differentiating among assets, e.g., whether they are tangible or intangible, financial or real, liquid or illiquid, etc. These are not defining characteristics of money, however.

omists is why do people hold an asset that is valueless except in exchange? The answer suggested here is that money is the only asset that provides exchange services that other assets cannot provide.[28]

A better question is why have all monetary economies evolved into fiat money economies? The answer to this question arises naturally from the framework presented here. In the analysis presented above, I implicitly assumed that commodity money is costless to produce and/or to maintain. This is not the case. Commodity monies require that resources be used in their production and to maintain the stock. In the case of commodity monies like precious metals, most of the costs are production costs (the maintenance costs, i.e., depreciation, are fairly low). In the case of more abstract commodity monies, like checkable deposits, the costs are on going, and are related to the extent of their use.[29] Indeed, even paper currency is not costless to produce and maintain.[30]

Because the production/maintenance of money requires the use of economic resources, the welfare gains associated with the use of resource-using money are necessarily smaller than if money were costless. If money production requires resources, augmenting the stock means that resources will be drawn, at least temporarily, from other uses. The nominal money stock (e.g., tons of gold or silver) will increase as long as the marginal exchange value of the last unit produced exceeds its marginal production cost. If there are maintenance costs, (e.g., the rate of physical depreciation is positive), then resource-using money will be less welfare enhancing because of higher maintenance costs. Because the verification costs are likely to be higher the more rapidly and less predictably an asset depreciates, there is an incentive to choose as money commodities that have a low, perhaps negligible, rate of depreciation. That is, all other things being the same, the asset with the lowest maintenance cost will serve as money.

For these reasons, society has an incentive to replace high-cost commodity money with lower cost money. One step in this evolution was to replace commodity money with lower cost *representative money*. A further step is to replace representative money with even lower cost fiat money. Because fiat money requires fewer resources for production and maintenance, its use is welfare enhancing. In addition, if fiat money were to further reduce transactions costs, social welfare could be further enhanced if specialization and trade were encouraged.

Despite its advantages over resource-using money, fiat money evolved slowly over a considerable period of time. Money's evolution was undoubtedly affected by wars, other political events, and difficulties associated with regulating the supply of various commodity monies. Nevertheless, the fact that fiat money increases society's welfare relative to commodity money suggests the evolution to fiat money is the result of economic forces rather than the happenstance of a number of noneconomic events (Russell, 1991).

Governments issue fiat money because private fiat money issuers have an incentive to issue money as long as the marginal value of the last nominal unit issued is greater than its production cost. Consequently, it would be difficult for private money issuers to make a credible commitment not to over issue fiat money, so as to make it "worthless" (e.g., Ritter, 1995).[31] Moreover, only the government can credibly commit to distribute the seigniorage revenue from money's creation. Seigniorage arises when the exchange value of the money issued exceeds money's production cost.

The existence of a money whose value in exchange exceeds its production cost has given rise to the notion that society's wealth exceeds the stock of tangible assets by the real value of fiat money held. As Tobin (1965, p. 676) put it:

> ...as viewed by the inhabitants of the nation individually, wealth exceeds the tangible capital stock by the size of what we might call the fiduciary issue. This is an illusion, but only one of the many fallacies of composition which are basic to any economy or any society.

Many economists accept Tobin's claim that the "wealth" associated with the real stock of fiat money is illusionary, but this proposition is erroneous. Tobin's error comes from viewing a monetary econ-

[28]See Tobin (1992, p. 774) for the traditional answer to this question.

[29]Thornton (1983) shows that the relevant issue for determining whether "inside money" is part of a society's stock of net wealth is whether there are resource costs involved in its production and maintenance.

[30]The United States is issuing another in a series of dollar coins. The purpose of these coins is to reduce the cost of maintaining the stock of currency, since coins depreciate less rapidly than paper money.

[31]Goodhart (1998) suggests that gold's role as a medium of exchange was greatly enhanced by government's use of gold to pay tributes or tariffs to avoid feuds.

omy as simply a barter economy with money. In so doing, he fails to recognize the private and public benefits that accrue from money's use. The benefits from the use of money naturally accrue to fiat money when society shifts from using a more costly commodity money to a less costly (or, ideally, completely costless) fiat money. Hypothetically, if resource-using money were replaced unit-for-unit with fiat money, the real value of the stock of fiat money would reflect the welfare benefits associated with the previously held stock of commodity money. Hence, the benefits of commodity money are embodied fully in the same real quantity of fiat money.[32]

Furthermore, the fact that all of the benefits from the previous stock of money would be obtained at lower cost guarantees that welfare is enhanced by the switch, even if there is no further reduction in the marginal transaction cost and, consequently, no further increase in trade and specialization.

Because of the existence of positive externalities associated with money's use, it is inappropriate to equate the welfare benefits of money with the real value of the money stock, i.e., M/P, as is frequently done. Nevertheless, it is clear that including the real value of the "fiduciary issue" as part of society's net wealth is not an illusion. Some time ago, Clower (1967) pointed out the dangers of treating monetary economies as if they were analytically equivalent to barter economies. In a similar vein, Coase (1960) argues that when a comparison of economies with alternative social arrangements is made, it is essential to consider the total effect. The true benefits of fiat money only can be obtained by comparing a fiat money economy with a commodity money economy or with barter or credit economies. The conclusion that the wealth associated with the real quantity of fiat money is illusionary emerges from a naive comparison of a fiat money economy with an economy where all of a sudden no one uses or holds money but nothing else changes.

Money and Credit as Media of Exchange

The world is dominated by monetary economies; however, this does not mean that transactions are not carried out using barter or credit. In monetary economies, all three methods of effecting exchange are used. Indeed, money may not be used to initiate most transactions. For example, when one considers every extension of trade credit or the transfer of goods by credit card, it is arguably the case that more transactions are carried out initially with credit than with cash or checks. Given the large and increasing use of credit in effecting transactions, how can one reasonably argue that this is a *monetary economy* and not a *credit economy*?

Let us begin this discussion by trying to answer the intriguing question: Could there be a pure credit economy with no medium of exchange? A pure credit economy may have been what Brunner and Meltzer (1971) had in mind when they argued that the problem of synchronizing payments and receipts could be achieved by making verbal promises. To see what such a world might look like, I will assume that not only is there perfect information, but that all individuals' promises are fully credible, i.e., no person makes a promise that cannot be kept.[33] In such a world, individual *A* could give individual *B* commodity *j* in exchange for a promise to receive commodity *j* or some other commodity at a later date. This world would be very complicated. For example, assume that individual *A* sells his labor services to Firm *F* for the promise from *F* to pay a certain quantity of commodity *j* at week's end. Individual *A* then buys the goods that he needs by promising to deliver *j* or some other commodity at some point in the future or by transferring part of Firm *F*'s promise to deliver commodity *j*. Of course, it is not necessary that these promises change hands *per se*, it could be that some centralized accountant keeps track of all promises made to and from all parties, or everyone could simply have a perfect memory.[34]

If promises were denominated in all possible commodities, quantities, and future dates, the problem of calculating the prices in this economy would be extremely difficult. The pricing problem could be significantly reduced (and the accounting simplified) if individuals agree to denominate all credit

[32]This point was initially made by Johnson (1969), p. 38, who recognized that there were utility or output gains associated with the use of money.

[33]There are a number of similarities between the no-money world I am about to describe and that described by Fama (1980, 1983). Others who have suggested that transactions could be carried out without the use of money are Black (1970), and Greenfield and Yeager (1983). See McCallum (1985) and White (1984) for analyses of these models.

[34]Kocherlakota (1998) suggests that fiat money is "merely a physical way of maintaining this balance sheet." Hence, he suggests that money is merely memory. He even suggests that his approach "represents an advance over the usual justifications for the existence of money: Money is a store of value, money is a medium of exchange and/or money is a unit account...After all, money does not allow society to transfer resources over time. Money does not reduce the cost of transferring resources from one person to another."

contracts in the same commodity. This would give rise to this commodity being a medium of exchange, however. For example, if all credits are denominated in *m*, it must be the case that a credit instrument worth *z* units of commodity *m* today must trade for *z* units of *m* itself. This means that individuals with *m* could simply trade it for the commodities they desire just as well as they could trade credit instruments denominated in *m*. If a credit instrument denominated in *m* facilitates trade, then so too must commodity *m*—*m* would be money.[35]

It could be that *m* is bulky, like a barrel of oil, so that it could not circulate hand-to-hand.[36] The promises themselves would be inconvenient, however, because longer-term contracts would have to be discounted relative to shorter-term contracts. This difficulty could be overcome by issuing non-interest-bearing sight drafts denominated in the common unit of account, i.e., currency.[37] In this case, a credit economy would give rise to money.

It could be, however, that the commodity is completely abstract, like a quark. Hence, we would have a pure credit, nonmonetary, exchange economy where all credit contracts are denominated in a unit of account, whose only function is to determine the price level (Fama, 1983). People, however, only would accept promises denominated in something abstract or something that they did care to hold if they were certain that they would be able to exchange these promises for the commodities they desire. Hence, that would demand that credit contracts be denominated in things that they value or are certain that they would be easily converted into other commodities.

Note the similarity between the world I have just described and the one that exists today. Our money is called the *dollar*. Congress adopted the *dollar* (and the decimal system) as our unit of currency in 1785. Alexander Hamilton's coinage recommendation establishing the U.S. dollar as 270 grains, 11/12 fine of gold or 416 grains, 0.89242 fine of silver was not adopted until April 1792.[38] Because of the inconvenience of carrying gold or silver, sight drafts were issued in convenient denominations. These claims on the U.S. stocks of gold and silver circulated in lieu of the commodities themselves. Over the years the dollar has been redefined. U.S. currency now is just a claim on the same quantity of U.S. currency. That is, we now have a pure paper currency standard. People are willing to hold intrinsically useless pieces of paper and claims that are denominated in intrinsically useless pieces of paper because they are certain that other individuals will accept the same. Collectively, the people agree to maintain the paper's value by limiting its issuance and to share the seigniorage.[39]

The above analysis also reinforces why it is efficient to have credit contracts denominated in the same commodity, and better still if this commodity is money. Jevons (1875) termed this the *standard of value* of money. The point to emphasize is that money facilitates the use of credit just as it facilitates the trade of consumable commodities and tangible assets, (e.g., savings deposits are exchanged for dollars that are used to purchase bonds). Consequently, while credit figures prominently in many transactions, the analysis presented above makes it clear why credit almost never is used sequentially for other transactions and why the adoption of a commodity medium of exchange has tended to precede credit arrangements, and not the other way around.[40]

Can credit instruments function as money? The answer is yes. They can and they have. Checkable deposits (or electronic transfers of funds) are the liabilities of the entities who hold the balances.[41] As such, they are promises to pay dollars upon demand. Such balances are included in measures of transactions money not only because they facilitate exchange, but also because financial institutions are committed to exchanging these deposits for cash immediately and at a fixed one-to-one ratio. This is what Pesek and Saving (1967) termed the

---

[35]Hoover (1988) has made this point in a similar fashion to argue against Fama's (1980) "new monetary economics."

[36]Fama (1980, 1983) eliminated the possibility that what he termed the "numeraire—unit of account" would circulate as money by assuming that it was a "barrel of oil." Note that this was not a necessary consequence of his model, but assumed. Hence, Fama did not establish that there would be no commodity that would circulate as a medium of exchange, rather he assumed it.

[37]In addition, there may be a problem with the denominations of such contracts. Indeed, Russell (1991) notes that both of these problems were drawbacks to bills of exchange circulating as currency during the seventeenth and eighteenth centuries in England.

[38]The mint began to coin silver in October 1794 and gold in July 1795, but a mistake by the first mint director resulted in coins of 9/10 fine. See Studenski and Krooss (1952) for more details.

[39]See McCallum (1985) for other ways of achieving price-level determinacy under a currency standard.

[40]Bagehot (1873) makes this point with respect to the origins of bank credit.

[41]See Goodfriend (1991) for a good discussion of the evolution of bank money.

instant repurchase clause. As long as the commitment is fully credible, such deposits and currency substitute perfectly. In this case, it is sensible to add such commitments to the stock of cash and call the sum the stock of money. Indeed, this is what is done.

Finally, it is worth noting that the increased use and availability of credit might mitigate the effects of disruptions to the supply of money, at least in the short run. In economies where the credit market is not well developed, a negative shock to the money supply may have a more immediate effect on output and/or prices than in an economy where individuals and businesses can not only readily borrow against their future income but can make transactions without having money immediately available. In addition, as more transactions are initiated with the use of credit, the stock of money necessary to support a given level of commodity transactions could diminish, i.e., the velocity of money could rise. It should be remembered, however, that financial transactions also require the use of money. In any event, it is reasonable to speculate that the relationship between money and output and money and prices is likely to change as financial markets develop and mature.

The Asset Demand for Money

The asset demand for money has been associated with two literatures. The first deals with demand for money as an asset and focuses on the interest elasticity of the demand for money. The second focuses on whether money should be defined to include non-medium-of-exchange assets. Money's essential function is to facilitate transactions. Hence, while it is appropriate to consider the effect of close substitutes for money on its demand, it is inappropriate to define money to include such non-medium-of-exchange assets.[42]

The asset demand for money focused attention on holding money for asset purposes, just like you hold any other asset. I will argue, however, that the asset demand for money is inconsequential.[43] The asset demand for money has its origins with Lavington (1968), but was most influentially advanced by Keynes. If money was held primarily as an asset, its demand should be quite sensitive to changes in interest rates, because the nominal return to holding money is zero.

If money is primarily a medium of exchange, however, the interest elasticity of money demand might be quite low. To see why, I note that Brunner and Meltzer (1971) begin their seminal work on money by noting that money remains in circulation even during periods of high and accelerating inflation. They argue that this fact "calls into question the relevance of treating money as an asset that provides little or no return."[44] The analysis of why individuals continue to use money during periods of high and accelerating inflation presented here is complementary with theirs. Money continues to function as a medium of exchange even under conditions of severe or hyperinflation because it enjoys a significant cost advantage over both barter and credit as a medium of exchange. Indeed, this advantage is likely to be so large that it would take an extreme increase in the holding cost to induce individuals to shift to the widespread use of either barter or credit to facilitate exchange. Moreover, the cost advantage of money increases as economies become increasingly specialized and dependent on exchange. The advantage also increases as payments practices become increasingly institutionalized.[45]

The point is that a large discontinuity exists between money and the next best alternative for exchange. Economists normally think of continuous functions where small changes induce individuals to switch from one alternative to another. No such continuum of media of exchange exists, however. Money so dominates barter and credit as a medium of exchange that it continues to serve as a medium of exchange despite very large increases in the cost of holding it. Jevons (1875, p. 6) was aware of this, stating:

> ...even if the medium of exchange varied considerably in value, people would go on making their payments in terms of it, as if there had been no variation, some gaining at the expense of others.

[42] See Mason (1976) for an excellent critique of this approach to defining money.

[43] McCallum and Goodfriend (1987) also have suggested the asset demand for money should be relatively inconsequential, stating that money "will also serve as a store of value, of course, but may be of minor importance to the economy in that capacity." They do not elaborate on why this should be so, however.

[44] Brunner and Meltzer (1971, p. 784).

[45] Wallace (1983) has emphasized one of these institutional features; namely, the legal restriction that currency is *legal tender*. Overlapping generations models focus on the store of value function of money, i.e., money's function as an asset (e.g., McCallum, 1983; and McCallum and Goodfriend, 1987).

The discontinuity between money and other means of exchange suggests that the demand for the medium of exchange may be rather insensitive to changes in its holding cost.

The key observation is the degree of substitutability between money and other assets is that the substitution is *unidirectional*: While money is an asset (or group of assets) that provides a particular function that other assets do not provide, at times, money may be held for the same reasons that other assets are held—the asset that normally serves as money also is now being held as a store of wealth. It is never the case, however, that other assets are held for the reason that money is primarily held. This means that when rapid and accelerating inflation significantly increases the cost of holding money it will not be a simple matter for other assets to substitute for it, i.e., become money. The most individuals can do is to economize on their money holdings along the lines suggested by Baumol (1952) and Tobin (1956). Since other assets dominate money in their ability to transfer wealth through time, however, individuals have a strong incentive to economize on their holdings of money for transactions purposes even when the returns to other stores of wealth are low.

Nevertheless, it is possible to envision circumstances where the return on real assets is so low that some individuals choose to hold money for the same reason they normally hold other assets. Indeed, classical economists, including Keynes, were concerned about the consequences of hoarding money. Given the observed stickiness of prices, they argued that hoarding money would have significant consequences for the real economy.

Hoarding money by individuals seems more likely, however, in economies with relatively poorly developed financial markets.[46] If few alternatives to holding wealth are readily available, more individuals may opt to hoard money, especially during times of economic or financial uncertainty. The more sophisticated and well developed the financial system becomes, however, the less likely it is that individuals will choose to hold money as an asset, even when nominal interest rates are extremely low.[47]

Keynes' notion of the asset demand for money focused the attention of monetary economists on the interest sensitivity of money demand. The interest sensitivity of money demand has been extensively investigated, with a wide array of results (e.g., Goldfeld and Sichel, 1990; and Laidler, 1993).

The amount of money held for transactions purposes depends on the planned volume of transactions. This, in turn, depends on the timing of receipts and payments, which are affected by the degree of specialization and the structure of the markets, as well as the size, extent, and activity in credit markets, etc. Changes in the opportunity cost of holding money will induce individuals to economize on their holdings of money balances, but the degree to which they do this depends on the size of the gain relative to the marginal cost of the economizing activity. Given that money holdings are typically a small part of an individuals' wealth and that individuals have a strong incentive to minimize their holding of money at any nonzero nominal interest rate, it would not be surprising to find a relatively low interest responsiveness of money demand. Indeed, empirical investigations of currency demand (e.g., Hess, 1971; and Dotsey, 1988), which has a zero nominal return and is held primarily for transactions purposes, suggest that the interest elasticity of currency demand is zero.[48] Other mediums of exchange that pay an implicit or explicit interest may be held, in part, for the same reasons individuals hold other assets, so that the demand for them is likely to be more sensitive to changes in their relative holding cost.

The Buffer-Stock Notion of the Demand for Money

The idea that there is no close substitute for money as a medium of exchange is complementary with the buffer-stock notion of money demand. In the buffer-stock theory (Laidler, 1984, 1987), holdings of real balances substitute for costly information and uncertainty. Individuals absorb shocks to their real money holdings due to a shock to their nominal money balances. Over time, nominal money holdings are adjusted to a level more consistent with individuals' demand for real money balances, given the level of nominal interest rates, the level

[46] Unfortunately, Keynes attempted to rationalize hoarding at a time when financial markets were well developed. Hence, it was difficult to explain why individuals held money when there were assets that had all of the same risk characteristics of money but yielded a positive rate of return (Barro and Fischer, 1976).

[47] It is usually assumed that zero is a lower bound for the nominal interest rate because individuals could simply hold money that bears a zero nominal return. This analysis too ignores the costs of acquiring and storing money. See Thornton (1999).

[48] Furthermore, most studies find a remarkably low substitutability between currency and transactions deposits, suggesting that these alternative media of exchange are held for quite different reasons.

and pattern of current income and expenditures and expectations of future nominal interest rates, income and expenditures, etc. The buffer-stock notion implies that individuals will not change their holdings of real money balances immediately when nominal interest rates, real income, or prices change.

Because other assets cannot perform money's function as a medium of exchange, I speculate that individuals respond more quickly to reductions in the real money balances due to negative nominal money shocks (or positive price-level shocks) than they do to increases caused by positive money shocks (or negative price-level shocks). For example, when there is a positive aggregate nominal money shock, individuals may hold these balances temporarily rather than spending them for goods and services or purchasing other assets. If this were to happen, there would be no immediate adjustment of output, employment, prices, or interest rates. On the other hand, since individuals cannot substitute for money, negative aggregate shocks may affect economic behavior directly and more quickly.

The Welfare Benefits of Money

The usual approach to assessing the welfare benefits of money is to assume that money is like other assets; for instance, shoes or cars. In the case of these assets, the benefits accrue only to the consumer so the welfare gains can be obtained by simply summing up the so-called *Harberger triangles*. It is well known that this approach fails when there are significant social externalities. Since I have argued that there are significant social benefits from money—because of the role it plays in expanding the size and extent of the markets for goods and credit, and the degree of specialization—this approach cannot possibly work. Indeed, it seems reasonable to speculate that the social benefits of money could eclipse its private benefits.

Unlike many innovations, it is virtually impossible to internalize the benefits from using money. This further enhances the idea that there is significant social benefit to money. Indeed, once the usefulness of money is recognized, the one who recognizes it has an incentive to share the insight with others, as my *parable of the trader* illustrates.

The Parable of the Trader

There was a producer who once every period loaded some of his produce on a wagon and went to a destination where he and other producers would meet to trade their wares. One day, the producer noticed that there was one good, *g*, that nearly everyone wanted and would exchange goods for *g*. Realizing that he can buy virtually any good he desired using *g*, he offers to take *g* for the goods he was trying to trade. Initially he does this only when the double coincidence of wants necessary for barter is lacking. He soon discovers, however, that trading in *g* is much faster and easier than searching out barter opportunities, so he stops seeking barter opportunities and his barter transactions become increasingly infrequent. By trading his wares for *g*, and *g* for the goods that he desires, this producer discovers that he can accomplish the desired trading in a fraction of the time that he had previously spent.

Now he could attempt to internalize the gain from his private knowledge (no one else has made this observation yet) by offering to tell others how they could save trading time for a fee. He realizes, however, that no one would pay for this information because all they have to do is observe him and they, too, would know the secret. More important, he realizes that he could further shorten his trading time if the others behaved as he. Hence, rather than keeping this information private and attempting to internalize the benefit from his superior information, the trader has an incentive to make the information public. In so doing, however, not only does he gain by shortening the transactions time, but others do as well.

As Laidler (1990, p. 48) puts it, "one agent's cash balances produce services not just for that agent then but for all other agents with whom his market activities bring him into contact." The use of money that facilitates the trade of one agent facilitates the trade of all agents. In addition, the reduction in individuals' exchange cost associated with money's use causes markets to flourish. Increased trade promotes greater specialization, greater dependence on trade, and a greater need for and use of money, and so on, and so forth.

The synergy among trade, money, and specialization makes isolating the welfare benefits of

money extremely difficult, if not futile. The welfare benefits of money can be ascertained only by comparing monetary economies with economies that have alternative arrangements for exchange, i.e., only by comparing the total welfare of a monetary economy with that of a nonmonetary economy.[49]

The Welfare Costs of Inflation

The main implication of the discontinuity between money and barter or credit as a medium of exchange is that money will continue to be used even at very high rates of inflation. This implies that the welfare costs of inflation, which are associated with the reduced reliance on money as a medium of exchange, may be relatively small. This is particularly likely at relatively modest rates of inflation. Hence, it is not surprising that estimates suggest that the cost of inflation is large only at relatively high inflation rates (e.g., Bruno and Easterly, 1996).

Furthermore, not only is it inappropriate to estimate the welfare gains from the use of money by adding up Harberger triangles, it is equally inappropriate to measure the welfare costs of inflation this way, as is frequently done.[50] Since money will continue to circulate as a medium of exchange and since the ability to economize further on money holdings is likely to be small, so, too, is the cost of inflation from holding money balances. This is important because many discussions about inflation assume that its principal cost is the private *shoe leather* cost associated with economizing on the use of money as a medium of exchange. If the externalities associated with money are important and significant, such analyses understate the welfare costs of inflation, perhaps significantly.

Most economists would argue that if an economy were just starting, the optimal rate of inflation would be zero. Nevertheless, many argue that once inflation is underway, society is better off tolerating some inflation rather than to suffer the output loss they believe would be associated with reducing inflation to zero. This idea is called Howitt's (1990) Rule. The effects of inflation on the institutional arrangements of trade are likely to be extremely important, however, and these costs are missed completely by estimates that ignore the externalities associated with money's roles as a medium of exchange and a standard of value. Consequently, Tobin's often cited dictum that "it takes a heap of Harberger triangles to fill an Okun gap," which underlies such analyses, is simply irrelevant if there are significant social costs of inflation.[51]

The third consequence of the discontinuity between money and other methods of exchange is that it may be inflation uncertainty, rather than inflation *per se*, that produces the most significant welfare cost. Here it is important to distinguish between the *medium of exchange* and *standard of value* functions of money. An important benefit of money is that it reduces shopping costs—gathering information about relative prices, planning, budgeting, etc. Uncertainty interferes with the shopping function by distorting price signals that enhance market efficiency. Price-level uncertainty makes distinguishing between absolute and relative prices and between permanent or transitory changes in the price level difficult. Distortions to the pricing mechanism affect the efficiency of markets that affect investment (e.g., DeLong and Summers, 1991; and Barro, 1995), financial markets, and relative input prices (e.g., Easterly, 1993). Inflation also reduces efficiency by encouraging the development of alternative market structures that would not exist in a world with a stable price level. Because uncertainty about the future level of prices increases with the average rate of inflation, these costs are likely to be small at relatively low rates of inflation but increase with the rate of inflation.

It could be, however, that the most deleterious effects of inflation on economic welfare may come from the effect of inflation on the efficient function of the credit market. Both the rate of inflation and

[49] This may have implications for how money is modeled. For example, it is frequently the case that money is modeled in the context of one good economy where exchange is implied but not explicitly modeled. Given the possibility that there are large externalities associated with money, this practice may not be useful for some issues. It also may have implications for other models. For example, Lucas (1980, p. 145) states, "When we apply theories of barter economies to problems in, say, public finance or labor economics, it is not our intent to obtain results applicable only to primitive or prehistoric societies. We apply this body of theory to money-using economies such as our own because we believe that for many problems the fact that money is used in attaining equilibrium can be abstracted from, or that the theoretical barter economy is a tractable, idealized model which approximates well (is well-approximated by) the actual monetary economy. If this practice is sound, then we want monetary theories which rationalize it or at least do not radically conflict with it."

[50] For example, Bailey (1956), Friedman (1969) and Lucas (1994).

[51] For a critique of some other limitations of Howitt's Rule, see Thornton (1996). Also see Marty and Thornton (1995) for a discussion of some other arguments for the desirability of moderate inflation.

inflation uncertainty are detrimental to denominating credit contracts in terms of fixed units of money. Consequently, while high, accelerating, and especially uncertain inflation may have a relatively small effect on money's medium-of-exchange function, they may have a significant effect on financial markets. It is not easy to replace money as the standard of value. Recently, credit contracts have been denominated in variable units of money, so that the value of the contract varies with a measure of the actual inflation experience during periods of inflation uncertainty. For reasons that are not well understood, however, this practice has been relatively limited, especially at relatively moderate inflation rates. Long-term debt markets tend to dry up during periods of rapid inflation and, as a consequence, the rate of capital formation slows. While far from definitive, the evidence suggests that the covariance between inflation and the rate of economic growth is negative (e.g., Bruno and Easterly, 1996).

While inflation potentially has a significant effect on the rate of economic growth, its potential to affect the level of output may be modest. To the extent that high and accelerating inflation reduces the reliance on money as either a medium of exchange or a standard of value, resources are drawn from one use to another. The result is that the level of measured output may change relatively little between high and low inflation states, but the distribution of output may be significantly different and the level of economic welfare may be significantly lower in higher inflation environments. This may account for the fact that economists have not found a statistically significant relationship between the rate of inflation and the level of output, at least for relatively moderate rates of inflation.

SUMMARY AND CONCLUSION

I have argued that money is a social arrangement resulting from a complicated evolutionary process. Money exists because it facilitates exchange by reducing the cost of trade. Seen in this point of view, money is but one of several institutional arrangements designed to reduce the costs of exchange. By reducing the cost of exchange, money reduces the reservation relative price where trade is mutually advantageous thereby encouraging more trade and greater specialization. Because of their strategic complementary, it is not surprising that money, trade, and specialization have tended to evolve simultaneously.

I argue that there are only three methods of effecting trade: simple and sequential barter, simple and sequential credit, and money. I then explain why the information and shopping costs of sequential barter and/or sequential credit are likely to be high relative to those of money. It is not surprising that the world is populated with monetary economies and not barter or credit economies.

I also have argued that money has a significant cost advantage relative to simple barter and credit and this advantage helps explain why the same good has served most often as both the medium of exchange and the unit of account, and why the development and widespread use of money tends to make simple barter scarce.

The use of money promotes specialization and trade by reducing exchange costs. The reduction in exchange costs associated with money cannot benefit one individual without benefiting others. Indeed, it is virtually impossible to internalize the benefits from money. Consequently, there are significant externalities associated with the use of money. Money is a social arrangement whose benefits can be calculated correctly only by comparing monetary economies with barter or credit economies. I speculate that the social gains from the use of money are likely to be large relative to the private opportunity cost of holding it. Furthermore, these benefits extend to nonresource-using fiat money. Indeed, the fact that nonresource-using money frees resources for production and/or leisure necessarily implies that, other things being the same, the transition from commodity to fiat money is welfare enhancing.

I argue that money enjoys an enormous cost advantage over barter or credit as a medium of exchange. Because of this, inflation is not likely to result in a large-scale substitution away from money as a medium exchange. Hence, money continues to circulate as a medium of exchange even during periods of hyperinflation. Significant costs of inflation could be associated with the effects of inflation uncertainty on the efficiency of the goods, labor, and financial markets, most especially the efficiency of the credit market because of the deterioration of money's function as a standard of value.

The fact that there are significant externalities associated with the use of money and that inflation increases the costs of using money gives rise to the possibility that the welfare costs of inflation are significant. Because money dominates barter and credit as a medium of exchange, the welfare costs of inflation due to a reduction in money's role as a medium

of exchange are likely to be small, relative to those associated with its function as a standard of value.

REFERENCES:

Alchian, Armen A. "Why Money?" *Journal of Money, Credit and Banking* (February 1977), pp. 133-40.

Archibald, Glen C., and Robert G. Libsey. "Monetary and Value Theory: A Critique of Lange and Patinkin," *The Review of Economic Studies* (October 1958), pp. 1-22.

Bagehot, Walter. *Lombard Street*, John Murray, London, 1873.

Bailey, Martin J. "The Welfare Cost of Inflationary Finance," *Journal of Political Economy* (April 1956), pp. 93-110.

Barro, Robert J. "Inflation and Economic Growth," Bank of England *Quarterly Bulletin* (May 1995), pp. 166-76.

_____ and Stanley Fischer. "Recent Developments in Monetary Theory," *Journal of Monetary Economics* (April 1976), pp. 133-67.

Baumol, William J. "The Transactions Demand for Cash: An Inventory Theoretic Approach," *Quarterly Journal of Economics* (November 1952), pp. 545-56.

Benhabib, J., and C. Bull. "The Optimal Quantity of Money: A Formal Treatment," *International Economic Review* (February 1983), pp. 101-11.

Black, Fischer. "Banking and Interest Rate in a World Without Money: The Effects of Uncontrolled Banking," *Journal of Bank Research* (Autumn 1970), pp. 9-20.

Braun R. Anton. "Another Attempt to Quantify the Benefits of Reducing Inflation," Federal Reserve Bank of Minneapolis *Quarterly Review* (Fall 1994), pp. 17-25.

Brunner, Carl, and Allan H. Meltzer. "The Uses of Money: Money in the Theory of an Exchange Economy," *American Economic Review* (December 1971), pp. 784-805.

Bruno, Michael, and William Easterly. "Inflation and Growth: In Search of a Stable Relationship," this *Review* (May/June 1996), pp. 139-46.

Bullard, James. "Testing Long-Run Neutrality Propositions: Lessons from the Recent Research," this *Review* (November/December 1999) pp. 57-77.

Clower, Robert W. "A Reconsideration of the Microfoundations of Monetary Theory," *Western Economic Journal* (December 1967), pp. 1-8.

Coase, Ronald H. "The Problem of Social Cost," *The Journal of Law & Economics* (October 1960), pp. 1-44.

Corbae, Dean, and Joseph A. Ritter. "Money and Search with Enduring Relationships," unpublished manuscript, 1998.

DeLong, J.B., and Lawrence H. Summers. "Equipment Investment and Economic Growth," *The Quarterly Journal of Economics* (May 1991), pp. 445-502.

Dornbusch, R., and Jacob Frenkel. "Inflation and Growth: Alternative Approaches," *Journal of Money, Credit and Banking* (February 1973), pp. 141-56.

Dotsey, Michael. "The Demand for Currency in the United States," *Journal of Money, Credit and Banking* (February 1988), pp. 22-40.

Duffy, John. "Monetary Theory in the Laboratory," this *Review* (September/October 1998), pp. 9-26.

Easterly, William. "How Much Do Distortions Affect Growth?" *Journal of Monetary Economics* (November 1993), pp. 187-212.

Fama, Eugene F. "Banking in the Theory of Finance," *Journal of Monetary Economics* (January 1980), pp. 39-57.

_____. "Financial Intermediation and Price Level Control," *Journal of Monetary Economics* (July 1983), pp. 7-28.

Fischer, Stanley. "Monetary Rules and Commodity Money Schemes under Uncertainty," *Journal of Monetary Economics* (January 1986), pp. 21-35.

Friedman, Milton. "The Optimum Quantity of Money," in *The Optimum Quantity of Money and Other Essays*, Aldine Publishing Co., 1969.

Goldfeld, Stephen M., and Daniel E. Sichel. "The Demand for Money," in the *Handbook of Monetary Economics*, Volume I, Benjamin M. Friedman and Frank H Hahn, eds., Elsevier Science Publishers, 1990, pp. 299-356.

Goodfriend, Marvin. "Money, Credit, Banking and Payments System Policy," Federal Reserve Bank of Richmond *Economic Review* (January/February 1991), pp. 7-23.

Goodhart, Charles A. E. "The Two Concepts of Money: Implications for the Analysis of Optimal Currency Areas," *European Journal of Political Economy* (1998), pp. 407-32.

Greenfield, Robert L., and Leland B. Yeager. "A Laissez Faire Approach to Monetary Stability," *Journal of Money, Credit and Banking* (August 1983), pp. 302-15.

Hess, Alan C. "An Explanation of Short-Run Fluctuations in the Ratio of Currency to Demand Deposits," *Journal of Money, Credit and Banking* (August 1971), pp. 666-79.

Hicks, John. *A Market Theory of Money*, Clarendon Press, Oxford, 1989.

Hirshleifer, J. "Exchange Theory: The Missing Chapter," *Western Economic Journal* (June 1973), pp. 129-46.

Hoover, Kevin D. "Money, Prices and Finance in the New Monetary Economics," *Oxford Economic Papers* (March 1988), pp. 150-67.

_____. "Some Suggestions for Complicating the Theory of Money," unpublished manuscript, 1995.

Howitt, Peter. "Zero Inflation as a Long-Term Target for Monetary Policy," in *Zero Inflation: The Goal of Price Stability*, Richard G. Lipsey, ed., C. D. Howe Institute, 1990.

Jevons, William Stanley. *Money and the Mechanism of Exchange*, Twentieth Century Press, 1875.

Johnson, Harry G. "Inside Money, Outside Money, Income, Wealth, and Welfare in Monetary Theory," *Journal of Money, Credit and Banking* (February 1969), pp. 30-45.

Johri, Alok. "On the Real Effects of Fiat Money in a Search Model," unpublished manuscript, Boston University, 1994.

Jones, R. "The Origin and Development of Media of Exchange," *Journal of Political Economy* (August 1976), pp. 757-75.

Kiyotaki, Nobuhiro, and Randall Wright. "A Search-Theoretic Approach to Monetary Economics," *American Economic Review* (March 1993), pp. 63-77.

_____. "A Contribution to the Pure Theory of Money," *Journal of Economic Theory* (April 1991), pp. 215-35.

_____. "On Money as a Medium of Exchange," *Journal of Political Economy* (August 1989), pp. 927-54.

Kocherlakota, Narayanna R. "Money Is Memory," *Journal of Economic Theory* (August 1998), pp. 232-51.

Laidler, David. "The Buffer-Stock Notion in Monetary Economics," *Conference Proceeding, Supplement to the Economic Journal* (March 1984), pp. 326-34.

_____. "Buffer-Stock Money and the Transmission Mechanism," Federal Reserve Bank of Atlanta *Economic Review* (March/April 1987), pp. 11-23.

_____. *Taking Money Seriously*, Humel Hempstead: Philip Allan; Cambridge, Mass.: MIT press, 1990, pp. 1-23.

_____. *The Demand for Money*, Harper Collins College Publishers, 1993.

Lavington, F. *The English Capital Market*, Augustus M. Kelley, New York, 1968.

Lucas, Robert E. Jr. "On the Welfare Cost of Inflation," unpublished manuscript, University of Chicago, 1994.

_____. "Equilibrium in a Pure Currency Economy," in *Models of Monetary Economies*, John H. Kareken and Neil Wallace, eds., Federal Reserve Bank of Minneapolis, 1980, pp. 131-45.

Marty, Alvin L., and Daniel L. Thornton. "Is There a Case for 'Moderate' Inflation?" this *Review* (July/August 1995), pp. 27-37.

Mason, Will E. "The Empirical Definition of Money: A Critique," *Economic Inquiry* (December 1976), pp. 525-38.

McCallum, Bennett T. "The Role of Overlapping Generations Models in Monetary Economics," *Carnegie-Rochester Conference Series on Public Policy* (Spring 1983a), pp. 9-44.

_____. "A Model of Commodity Money: Comments," *Journal of Monetary Economics* (July 1983b), pp. 189-96.

_____. "Bank Deregulation, Accounting Systems of Exchange, and the Unit of Account: A Critical Review," *Carnegie-Rochester Conference Series on Public Policy* (Autumn 1985), pp. 13-45.

McCallum, Bennett T., and Marvin S. Goodfriend. "Demand for Money: Theoretical Studies," *New Pargrave: A Dictionary of Economics*, John Eatwell, Murray Mulgate and Peter Newman, eds., MacMillan Publishing, 1987, pp. 775-80.

Meltzer, Allan H. "Information, Sticky Prices, and Macro-economic Foundations," this *Review* (May/June 1995), pp. 101-18.

Menger, Carl. "On the Origin of Money," *The Economic Journal* (June 1892).

Mundell, Robert. *Monetary Theory, Inflation, Interest, and Growth in the World Economy*, Goodyear Publishing Co., 1971.

Niehans, Jurg. "Money and Barter in General Equilibrium with Transaction Costs," *American Economic Review* (December 1971), pp. 773-83.

_____. *The Theory of Money*, Johns Hopkins University Press, 1978.

Oh, Seonghwan. "A Theory of a General Acceptable Medium of Exchange and Barter," *Journal of Monetary Economics* (January 1989), pp. 101-19.

Ostroy, Joseph M. "The Informational Efficiency of Monetary Exchange," *American Economic Review* (September 1973), pp. 597-610.

Patinkin, Don. *Money, Interest, and Prices: An Integration of Monetary and Value Theory*, Harper and Row, second edition, 1965.

Pesek, Boris P., and Thomas R. Saving. *Money, Wealth and Economic Activity*, MacMillan Company, 1967.

Ritter, Joseph A. "The Transition from Barter to Fiat Money," *American Economic Review* (March 1995), pp. 134-49.

Russell, Steven. "The U.S. Currency System: A Historical Perspective," this *Review* (September/October 1991), pp. 34-61.

Smith, Adam. *An Inquiry into the Nature and Causes of the Wealth of Nations*, Edwin Cannan, ed. Random House, 1937.

Studenski, Paul, and Herman E. Krooss. *Financial History of the United States*, McGraw-Hill Book Company, 1952.

Thornton, Daniel L. "Bank Money, Net Wealth and the Real Balance Effect," *Journal of Macroeconomics* (Winter 1983), pp. 105-17.

_____. "The Costs and Benefits of Price Stability: An Assessment of Howitt's Rule," this *Review* (March/April 1996), pp. 23-38.

_____. "Nominal Interest Rates: Less Than Zero?" *Monetary Trends* (January 1999) p. 1.

Tobin, James. "Money," in *The New Palgrave Dictionary of Money and Finance*, by Peter Newman, Murray Milgate and John Eatwell, eds., MacMillian Press, 1992, pp. 770-79.

_____. "Money and Economic Growth," *Econometrica* (October 1965), pp. 671-84.

_____. "The Interest-Elasticity of the Transactions Demand for Cash," *Review of Economics and Statistics* (August 1956), pp. 241-47.

Trejos, Alberto, and Randall Wright. "Search, Bargaining, Money, and Prices: Recent Results and Policy Implications," *Journal of Money, Credit and Banking* (August 1993, Part 2), pp. 558-76.

Wallace, Neil. "A Legal Restrictions Theory of the Demand for 'Money' and the Role of Monetary Policy," Federal Reserve Bank of Minneapolis *Quarterly Review* (Winter 1983), pp. 1-7.

White, Lawrence H. "Competitive Payments Systems and the Unit of Account," *American Economic Review* (September 1984), pp. 699-712.

Appendix

The Effect of Exchange Costs of the Exchange Ratio Necessary for Trade

Robert D. Dittmar and Daniel L. Thornton

The text argues that costly trade reduces the feasible range of exchange ratios where trade is mutually advantageous. The question that arises is by how much must the exchange ratio change to compensate an individual for the costs of exchange if there are fixed exchange costs? The question is not answered easily because the relevant comparison is the utility levels obtained with costly trade and with autarky. What must the terms of trade be to compensate an individual for fixed transactions costs? Some intuition about this can be obtained by considering the effect of a small change in fixed costs in the borderline case when the individual is indifferent between the autarkic consumption bundle and the consumption bundle obtained by paying a fixed cost and trading.

Let (c_1^A, c_2^A) denote the individual's consumption bundle under autarky, and (c_1^T, c_2^T) denote the individual's consumption bundle with fixed exchange costs. Indifference implies that

$$U(c_1^T, c_2^T) = U(c_1^A, c_2^A). \tag{A.1}$$

When production technologies are linear and the terms of trade, λ, are such that $\lambda \neq f_2'/f_1'$, individuals specialize in the production of one of the goods, good 1 or good 2. The budget constraint

$$\lambda c_1^T + c_2^T = \lambda f_1(\Gamma) - \Omega \tag{A.2}$$

is satisfied for individuals specializing in good 1, and

$$\lambda c_1^T + c_2^T = f_2(\Gamma) - \Omega \tag{A.3}$$

for individuals specializing in good 2.

The optimality of (c_1^T, c_2^T) with trade implies that the individual equates the ratio of the marginal utilities of the two goods to the exchange ratio, λ, so that

$$U_1(c_1^T, c_2^T) = \lambda U_2(c_1^T, c_2^T) \tag{A.4}$$

is satisfied.

Equation A.4, equation A.1, and one of the two budget constraints above implicitly determine the consumption bundle in the case of specialization and the terms of trade that are necessary to compensate the individual for trading when there is a fixed transaction cost, Ω.

In principle these equations can be solved to determine the effect of Ω on λ when an individual specializes in the production of either of the two goods. A closed-form solution cannot be obtained, however, without making explicit assumptions about functional forms. Linear approximations to these functions that will be accurate predictors of the effects of small transaction costs can be made, however.

These linear approximations are obtained by implicitly differentiating the equations and evaluating the resulting expressions at $\Omega = 0$. Note that if $\Omega = 0$, $\lambda = f_2'/f_1'$ if an individual is to be indifferent between trading and autarky. Furthermore, at these terms of trade, the individual must be indifferent between specializing in the production of good 1 or good 2. Consequently, either of the budget constraints above can be used as the starting point of the approximation. Finally, note that optimization requires the individual to equate the ratio of the marginal utilities of the two goods to terms of trade. In the absence of exchange costs, or under autarky, the condition is

$$\frac{U_1(c_1^A, c_2^A)}{U_2(c_1^A, c_2^A)} = \frac{f_2'}{f_1'}. \tag{A.5}$$

To economize on notation, we use an overdot to represent differentiation with respect to the fixed cost of exchange, Ω, i.e., $\dot{\lambda} = d\lambda/d\Omega$. Regardless of whether the individual specializes in good 1 or 2, we can implicitly differentiate Equation A.4 to obtain,

(A.6) $$U_1(c_1^T, c_2^T)\dot{c}_1^T + U_2(c_1^T, c_2^T)\dot{c}_2^T = 0.$$

Evaluating this expression at $\Omega = 0$ yields

(A.7) $$\frac{f_2'}{f_1'}\dot{c}_1^T(0) + \dot{c}_2^T(0) = 0.$$

Implicitly differentiating Equation A.2, on the assumption the individual specializes in good 1, yields

(A.8) $$\dot{\lambda}c_1^T + \lambda\dot{c}_1^T + \dot{c}_2^T = \dot{\lambda}f_1(\Gamma) - 1.$$

Evaluating this expression at $\Omega = 0$, yields

(A.9) $$\dot{\lambda}(0)c_1^A + \frac{f_2'}{f_1'}\dot{c}_1^T(0) + \dot{c}_2^T(0) = \dot{\lambda}(0)\, f_1(\Gamma) - 1.$$

Combining Equations A.7 and A.9 and solving for $\dot{\lambda}(0)$ yields

(A.10) $$\dot{\lambda}(0) = \frac{1}{f_1(\Gamma) - c_1^A}.$$

Equation A.10 shows the effect on λ of a small fixed transaction cost, evaluated at the point $\Omega = 0$. Hence, an individual will be induced to specialize in good 1 and trade only if the terms of trade are approximately

(A.11) $$\lambda(\Omega) \approx \frac{f_2'}{f_1'} + \left(\frac{1}{f_1(\Gamma) - c_1^A}\right)\Omega$$

or larger. This result is intuitive. The quantity $f_1(\Gamma) - c_1^A$ is approximately the amount of good 1 that the individual must give up to trade. Hence, $\Omega/(f_1(\Gamma) - c_1^A)$ is the extra amount of good 2, per unit of good 1 exchanged, that the individual must obtain to be compensated for the fixed cost of entering the market.

A similar analysis applies to individuals that specialize in the production of good 2. In this case, Equation A.3 is differentiated to obtain

(A.12) $$\dot{\lambda}c_1^T + \lambda\dot{c}_1^T + \dot{c}_2^T = -1.$$

Evaluating Equation A.12 at $\Omega = 0$, as before, yields,

(A.13) $$\dot{\lambda}(0)c_1^A + \frac{f_2'}{f_1'}\dot{c}_1^T(0) + \dot{c}_2^T(0) = -1.$$

Combining Equations A.7 and A.13, yields

(A.14) $$\dot{\lambda}(0) = -\frac{1}{c_1^A}.$$

Hence, an individual will be induced to specialize in good 2 and trade only if the terms of trade are

(A.15) $$\lambda(\Omega) \approx \frac{f_2'}{f_1'} - \frac{1}{c_1^A}\Omega$$

or smaller. The quantity $-\Omega/c_1^A$ is the discount per unit of good 1 purchased required to compensate the individual for the fixed cost of trade.