The U.S. trade deficit: Made in China?

Energy markets and the Midwest economy

Shopping without cash: The emergence of the e-purse

Consumption-based macroeconomic forecasting

Index for 2005
# Contents

Fourth Quarter 2005, Volume XXIX, Issue 4

<table>
<thead>
<tr>
<th>Page</th>
<th>Title</th>
<th>Authors/Contributors</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>The U.S. trade deficit: Made in China?</td>
<td>Chad P. Bown, Meredith A. Crowley, Rachel McCulloch, and Daisuke J. Nakajima</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rapid growth of the U.S. bilateral trade deficit with China has promoted a widespread view that the overall trade deficit is “made in China.” The authors examine the probable consequences of increased protection directed toward U.S. imports from China. Their appraisal of recent and prospective U.S. trade policy focuses on textiles and apparel—sectors where the growth of imports from China has been prominent. They also consider the likely effects of yuan appreciation on the bilateral and overall trade deficits.</td>
</tr>
<tr>
<td>19</td>
<td>Energy markets and the Midwest economy</td>
<td>Richard H. Mattoon</td>
</tr>
<tr>
<td></td>
<td></td>
<td>This article examines the effects of fuel prices on the Midwest economy. As a region that is still disproportionately reliant on manufacturing, the Midwest would be expected to feel the brunt of high energy costs. However, is this still the case, or has the impact been mitigated over time by the restructuring of the region’s economy? The author finds that the Midwest has followed the national pattern of decreasing reliance on energy to produce regional output, but still has a relatively high concentration of energy-intensive industries that, when combined with unfavorable weather conditions, suggests that the region is still sensitive to energy price spikes.</td>
</tr>
<tr>
<td>32</td>
<td>Conference on Bank Structure and Competition call for papers</td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>Shopping without cash: The emergence of the e-purse</td>
<td>Carol L. Clark</td>
</tr>
<tr>
<td></td>
<td></td>
<td>This article finds that successful e-purse programs tend to have a captive audience that drives critical mass, such as those found in the transportation industry or government sector; an affordable cost structure relative to other payment instruments; compelling incentives for consumers and merchants; and technology that is well tested and addresses standards issues before rollout.</td>
</tr>
<tr>
<td>52</td>
<td>Consumption-based macroeconomic forecasting</td>
<td>Jeffrey R. Campbell and Spencer D. Krane</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The authors build a small-scale econometric model based on the permanent income theory of consumption and balanced economic growth in order to study the influence of permanent and transitory factors on the level of economic activity.</td>
</tr>
<tr>
<td>71</td>
<td>Index for 2005</td>
<td></td>
</tr>
</tbody>
</table>
The U.S. trade deficit: Made in China?

Chad P. Bown, Meredith A. Crowley, Rachel McCulloch, and Daisuke J. Nakajima

**Introduction and summary**

Most economists and policymakers agree that open international markets create important mutual benefits for the United States and its trading partners. By encouraging countries to focus on the activities in which they have a comparative advantage, the free international exchange of goods and services allows all countries to raise their living standards. Accordingly, the United States played a central role in the establishment in 1995 of the World Trade Organization (WTO), which now sets the ground rules for national trade policies worldwide and helps resolve trade disputes among its members.1 Within the WTO framework, U.S. officials have worked to eliminate remaining tariffs, quotas, and other policy barriers to trade. At the same time, they have also negotiated trade flows with particular partners, such as the North American Free Trade Agreement (NAFTA) with Canada and Mexico.

Ongoing global trade liberalization has contributed to a major expansion of both exports and imports for the United States and most other countries. However, freeing trade from policy restrictions does not ensure that trade among nations will be balanced, either overall or with any particular trading partner. Aggregate trade deficits or surpluses reflect underlying macroeconomic conditions in each country. In some countries, such as China in recent years, exports have grown faster than imports, resulting in an overall trade surplus. In other countries, including the United States, imports have grown faster than exports, resulting in an overall trade deficit.2 Moreover, China has emerged as a major source of U.S. imports, leading to a widespread view that the record overall U.S. trade deficits of recent years are “made in China.” A large and persistent trade imbalance may raise policy concerns because of its perceived links to domestic production and employment—specifically, the fear that more imports will mean less production and fewer jobs in the United States. A large and persistent trade deficit may also be worrisome to the extent that it increases U.S. reliance on international borrowing—the sale abroad of U.S. bonds and other securities.

Overall U.S. performance in trade is most frequently reported and publicized as the trade balance, or more specifically as the merchandise trade balance. This number, released monthly by the U.S. Department of Commerce, is the difference between the dollar value of all U.S. exports of tangible goods (merchandise) and the dollar value of all U.S. imports of tangible goods. Broader measures include the balance on goods and services trade and the current account balance.3 Figure 1 illustrates the recent increase in the merchandise trade deficit as a share of gross domestic product (GDP). In 2004, the U.S. merchandise trade deficit topped $665 billion or 5.6 percent of GDP.4 This deficit set a new record, and most analysts are forecasting an even larger trade gap in 2005.

The overall trade balance can be expressed as the sum of the bilateral trade balances added up over all trading partners or, alternatively, as the sum of the industry-level trade balances added up over all industries. In a trading system with many countries and many goods, there is no theoretical reason to expect trade to be balanced with any particular partner or in any

---

**Note**

1. Chad P. Bown is an associate professor in the Department of Economics and International Business School, Brandeis University, and a nonresident fellow of the Brookings Institution. Meredith A. Crowley is an economist at the Federal Reserve Bank of Chicago. Rachel McCulloch is the Rosen Family Professor of International Finance at Brandeis University. Daisuke J. Nakajima is an associate economist at the Federal Reserve Bank of Chicago. The authors thank Stephen Cecchetti, Rashmi Shanikar, Blake LeBaron, Michael Koupastas, and Craig Furfine for helpful comments and suggestions.
that a reduction in the U.S. trade deficit with China will translate into a similar reduction in the nation’s overall trade deficit. The purpose of this article is to examine the bilateral trade relationship with China, particularly the likely consequences of instituting measures intended to limit U.S. imports from China. Our appraisal of recent and prospective U.S. trade policy focuses on textiles and apparel, sectors where the growth of imports from China has been especially prominent. We also consider the role that yuan appreciation might play in shrinking the bilateral and overall trade deficits.

We begin our analysis by examining the macroeconomic factors that shape the trade balance. The U.S. merchandise trade deficit is necessarily equal in size to the difference between the nation’s total domestic goods production and its total expenditure on goods. Sectoral trade policies, such as recent measures to limit certain apparel imports from China, can change the composition of U.S. trade by shifting U.S. import demand toward other foreign suppliers or substitute products. However, reduction of the overall trade deficit will occur only to the extent that total U.S. production rises or total U.S. expenditure falls; neither is a probable outcome of targeted trade policies affecting only selected trading partners and products. Yuan appreciation might likewise
change the composition of U.S. trade by shifting U.S. demand toward alternative suppliers. Again, the overall trade gap can change only to the extent that changes in the international value of the dollar result in more U.S. production or less U.S. expenditure. However, an appreciation of the yuan against the dollar that is not fully anticipated by market participants would likely be associated with an unexpected reduction in the rate of accumulation of U.S. dollar-denominated assets by the Chinese. This, in turn, could cause an unexpected increase in U.S. interest rates. Because higher interest rates tend to reduce U.S. expenditure, yuan appreciation could reduce the overall U.S. trade gap as well as the size of the bilateral deficit with China. Although it is too soon to evaluate the full consequences of China’s new exchange rate policy, one immediate effect of the small appreciation announced on July 21, 2005, has been to change expectations—most market participants now see further appreciation as likely.

In the second part of this article, we take a closer look at current and prospective U.S. trade policies toward China. In the 1980s, U.S. policymakers took steps to limit surging imports from Japan (see box 1). Now they seem inclined to take similar steps to reduce U.S. imports from China. Although sector-specific measures cannot have an important effect on the overall U.S. trade imbalance, they may have a significant impact on trade flows at the industry level, including U.S. and Chinese trade with other nations. Past experience with U.S. policies targeting specific foreign suppliers, such as Japan and Korea, suggests that restrictions on U.S. imports from China are more likely to divert U.S. demand toward other low-cost foreign suppliers than toward domestic import-competing industries (Moore, 1996; Prusa, 2001; Bown, 2004). Such restrictions may also affect the flow of Chinese exports toward other potential markets (Bown and Crowley, 2005b), perhaps causing other importing countries to impose their own trade barriers on the same types of products. Indeed, research suggests that the probable result of selective trade policies is the reduction of national and world well-being without a significant effect on the overall trade balance.

External imbalances and domestic macroeconomic fundamentals

We have noted that the overall trade balance can be expressed as the sum of the bilateral trade balances added up over all trading partners or as the sum of the industry-level trade balances added up over all industries. However, an alternative approach comparing domestic production of goods and domestic expenditure on goods reveals an important relationship between the overall trade balance and macroeconomic conditions. For each category of goods at the industry or subindustry level, U.S. net exports (exports minus imports) must equal total domestic production of such goods minus total domestic expenditure for such goods. Summing over all categories of goods, this implies that the nation’s merchandise trade balance must equal the difference between total goods production and total goods expenditure. The overall trade deficit thus reflects macroeconomic conditions and can change only to the extent that those macroeconomic conditions change. Although sectoral policies, such as tariffs and quotas, can have important effects on the trade balances of particular industries or with particular trading partners, these policies can reduce the overall trade deficit only to the extent that they affect macroeconomic conditions, that is, to the extent that they increase total domestic production or reduce total domestic expenditure.

The merchandise trade balance reflects only transactions involving tangible goods, which account for a declining share of total U.S. production and expenditure. To link the nation’s overall income and expenditure to its international position, we use the current account, which includes not only transactions involving tangible goods but also trade in services, net foreign income, and unilateral transfers. In macroeconomic terms, the current account deficit is equal to the difference between the nation’s total income and its expenditure on consumption and domestic investment or, equivalently, to the gap between the nation’s total (public and private) savings and its total domestic (public and private) investment spending.

The counterpart to the current account balance is the financial account balance, which records net U.S. spending for foreign assets (U.S. capital outflows) and foreign spending for U.S. assets (U.S. capital inflows). The U.S. financial account surplus, roughly equal in size but opposite in sign to the current account deficit, is loosely interpreted as indicating U.S. net borrowing from the rest of the world—the borrowing needed to cover the gap between U.S. spending and receipts as recorded in the current account. The large U.S. current account deficit and the large increase in U.S. borrowing from the rest of the world are therefore two sides of the same coin.

While reports in the popular press often portray the trade and current account deficits as alarming developments, economic theory shows that a trade or current account deficit (or surplus) is inherently neither bad nor good in itself. Whether a deficit should be seen as a problem depends on a country’s present and future circumstances. The deficit is a benign development only
FIGURE 3
Components of the U.S. current account, percent of GDP

![Chart showing components of the U.S. current account, percent of GDP.](chart.png)

Source: U.S. Bureau of Economic Analysis.

as long as the required level of borrowing is reasonable, given the country’s economic prospects and the interest rate on borrowed funds. Thus, policymakers may become concerned if the current account deficit grows sufficiently large.11

Although rapid growth of the merchandise trade deficit underlies recent growth of the U.S. current account deficit (see figure 3), the current account deficit and the trade deficit differ in terms of public response and the political fallout. The two main issues raised by the large and persistent U.S. current account deficit are sustainability and intergenerational equity. The sustainability issue concerns the prospect for financing an ever-growing current account deficit—how long lenders abroad will continue to supply capital to the United States on the same terms, and what will happen if and when they stop. Intergenerational equity concerns implications for the well-being of future U.S. citizens. A current account deficit (financial account surplus) means that a country is borrowing more from residents of other countries than it is lending to them. The nation’s increasing indebtedness to the rest of the world may be viewed as placing an unfair burden on future generations. In contrast, concerns about the trade deficit focus on its implications for overall macroeconomic performance, as well as for employment and output in particular “trade-sensitive” manufacturing industries, such as steel, autos, textiles, and apparel.

However, empirical analysis reveals no causal link from trade deficits to macroeconomic performance or job creation.12

**Why focus on China?**

Analysts estimate that the U.S. current account deficit, now 6 percent of GDP, would need to drop to 2–3 percent of GDP in order to achieve long-run sustainability (Kouparitsas, 2005; Roubini and Setser, 2004).13 Thus, without reductions in other bilateral imbalances, sustainability could not be achieved even if the U.S.–China trade deficit dropped to zero. Furthermore, unless accompanied by changes in overall macroeconomic conditions in the United States, any reduction in the U.S. deficit on trade with China, whether achieved through trade policy or yuan appreciation, would cause corresponding increases in other bilateral deficits, as U.S. demand for products previously purchased from China was diverted to other foreign suppliers. Thus, the effect on overall trade and current account balances would likely be minimal. So why emphasize the role of China?

To begin with, China currently has the largest bilateral trade surplus with the United States; relative to the size of China’s economy, the surplus is even larger.14 And China’s bilateral surplus is also among the fastest growing. To the extent U.S. policymakers believe that bilateral balances with other regions will not be affected, they may see policies aimed at China as the most efficient way to make progress in bringing down the overall deficits. Moreover, most researchers believe that China has kept the value of its currency artificially cheap, thus implicitly subsidizing its exports and taxing its imports.15 This view is bolstered by China’s large accumulation of official U.S. dollar reserves. China’s unwillingness to allow the yuan to appreciate has, in turn, made other Pacific Rim countries reluctant to allow their own currencies to appreciate, lest export sales be lost to Chinese rivals. Following China’s recent announcement of its new exchange rate regime, Malaysia responded by shifting its own currency regime from a dollar peg to a basket peg. However, given the very small initial change in the yuan’s value, most countries in the region have adopted a wait-and-see attitude.

At the sectoral level, many of China’s exports to the United States compete with domestic manufacturing industries already facing stiff competition from other...
Rapid export-led growth and escalating rhetoric: The basics of today’s U.S.–China confrontation over trade and currency misalignment closely resemble those of the U.S.–Japan trade conflict in the 1980s. Just as today, many U.S. officials seized upon the large and growing bilateral trade deficit with Japan as a “smoking gun”—seemingly incontrovertible evidence that America’s problems were rooted in trade and currency practices abroad (McCulloch, 1988). In some specifics, the two situations are strikingly similar. Just as today, the United States in the 1980s was running not only a large bilateral trade deficit with Japan but also large overall trade and current account deficits. Also similar to the recent situation with China, Japan in the 1980s prevented its currency, the yen, from appreciating relative to the U.S. dollar through massive official purchases of U.S. securities, thus allowing the United States to finance a large and growing U.S. fiscal deficit without driving up U.S. interest rates. Mann (2005) describes the recent U.S. relationship with China (and other surplus nations) as one of co-dependency; China’s willingness to add to its already huge stock of official U.S. dollar assets has provided the United States with cheap financing and thereby, to some extent, sustained the high level of U.S. expenditure.

The similarities extend beyond macroeconomic roots. In both Japan and China, government subsidies to export industries played at least a supporting role in export success. Moreover, both countries’ exports were subjected to country-specific U.S. trade policies—policies benefiting established U.S. trading partners as well as, and sometimes more than, competing domestic firms. Beginning in 1981, a voluntary restraint agreement limited Japan’s fast-growing exports of autos to the United States. Today bilateral actions limit China’s fast-growing exports of textiles and apparel to the U.S. market, and Chinese exporters in a number of other industries face high U.S. antidumping duties.

However, there are also notable differences between the two situations. In the 1980s, Japan’s overall imports, and especially imports of intermediate goods, were low compared with other industrialized economies. In contrast, China’s overall imports are large and growing; China’s goods production is highly integrated with the world economy. While Japan in the 1980s ran a sizeable overall trade surplus as well as a large bilateral surplus on trade with the United States, the rapid growth of China’s imports has meant that its overall trade surplus, though recently trending upward, is much smaller than its bilateral surplus on trade with the United States. Moreover, where multinational corporations had little success in penetrating Japan in the 1980s, China has become an important magnet for foreign direct investment and ranked first among developing host countries in 2004. By most measures, China can be considered a very open economy—even more so given its size and level of development.

Perhaps the most important difference is that Japan’s total labor force limited the growth of its economy and its exports. In the 1980s, Japan was already near full employment. In contrast, China is only in the early stages of mobilizing its huge labor supply. Many unemployed and underemployed workers, including those displaced through the restructuring of agriculture and state-owned enterprises, have yet to be drawn into the more market-oriented and efficient parts of the Chinese economy. Thus, China’s rapid growth has the potential to continue for decades. If the current growth rate were to be maintained, China could overtake the United States as the world’s largest economy—and most important market—within a few decades.

Do the consequences of the intense U.S. policy focus on Japan in the 1980s shed any light on the current situation with respect to China? The results of U.S. efforts included a substantial appreciation of the Japanese yen relative to the dollar beginning in February 1985 and bilateral trade agreements for automobiles, semiconductors, agriculture, and other sectors in which U.S. negotiators sought to restrict imports from Japan or increase U.S. exports to Japan. Although it is not possible to identify specific causal links to subsequent developments, some of these developments are worth noting in connection with recent U.S. policy initiatives toward China.

First, these measures did not close the bilateral trade gap. Japan continued to maintain a substantial surplus on trade with the United States as well as a global trade surplus despite bilateral trade measures and yen appreciation. As figure 2 (p. 3) shows, the U.S. bilateral imbalance on trade with Japan has only recently been eclipsed in size by the record deficit on trade with China. Yet, Japan’s continuing surpluses on trade did not translate into continuing vigorous growth of the Japanese economy. Perhaps as a consequence of macroeconomic policies aimed at countering the effects of the strong yen, Japan’s economic growth slowed to a crawl during the “lost decade” of the 1990s (Bergsten, Ito, and Noland, 2001); meanwhile, the United States experienced vigorous growth but also a rising trade deficit.

Effects at the sectoral level are also instructive. Faced with negotiated limits on the volume of U.S. auto imports from Japan, Japanese producers upgraded the quality and price of cars destined for the U.S. market (Feenstra, 1988). U.S. auto imports from Europe and later Korea surged, while many Japanese companies established U.S. production facilities. Japan’s Toyota now rivals General Motors and Ford in U.S. sales and far exceeds them in profitability. Autos and auto parts still dominate Japan’s exports to the United States.

Is China the new Japan?

BOX 1

1 Other things being equal, larger countries tend to have a lower ratio of trade to GDP; likewise, developing countries tend to have a lower ratio of trade to GDP.
foreign suppliers. Foremost among these industries are apparel and textiles, sectors with significant protection from competing imports. Although apparel and textiles together now account for a low share (9 percent in 2004) of total U.S. imports from China, and that share has been declining (see figure 4), China’s role at the industry level is significant and growing. China is currently the largest foreign supplier, accounting for about one-fifth of total U.S. imports. Also, Chinese exports of textile and apparel products to the United States have continued to grow as a share of total U.S. imports at the industry level (figures 5 and 6) and as a share of U.S. domestic consumption (figure 7).

As we document below, U.S. trade policy since the 1990s has imposed greater restrictions on imports from China than on those originating elsewhere. Moreover, as a new entrant to world markets, China has been at a disadvantage in contesting U.S. trade policies that limit its exports. Before its accession to the WTO in 2001, China could not bring trade disputes under WTO rules and had no formal role in multilateral negotiations. Even though China is now in the WTO, the terms of its accession allow for a long transitional period of potentially discriminatory treatment by other WTO members. Moreover, China has not been included in recent U.S. efforts to negotiate preferential trade agreements. Therefore, Chinese exports must compete in the U.S. market with goods from an increasing number of countries that face lower trade barriers.

U.S. officials may also be looking ahead to competition with China in other sectors. Despite the recent emphasis on textiles and apparel, where U.S. imports from China soared in early 2005 following the elimination of U.S. quotas on these products, Chinese competition in other manufacturing industries has been growing even more rapidly (see figure 5). As with textiles and apparel, to a large extent these Chinese gains have come at the cost of traditional exporters. The depreciation of the U.S. dollar relative to the euro and the Canadian dollar has reinforced this trend by redirecting U.S. import demand toward China, Japan, and other East Asian exporters. Moreover, Chinese producers have proved to be adept at producing increasingly sophisticated products. Chinese auto parts are already entering the U.S. market in substantial quantities; imports of Chinese-built vehicles are expected as early as 2007 (Power, 2005).
These factors have combined to produce growing support both from within the United States and from established U.S. trading partners for new action to limit China’s access to U.S. markets. So far, pressure on China has been focused primarily in two areas: exports of apparel and the yuan-dollar exchange rate. But industry-specific measures are unlikely to have a noticeable effect on the overall U.S. trade deficit. Reducing U.S. imports of apparel through measures targeted only at China would mainly divert U.S. import demand toward other foreign suppliers, thus restoring some of the sales these countries have recently lost to their Chinese competitors. Although overall U.S. well-being would likely decline as American consumers face higher prices, some benefits would be reaped by rival producers abroad and perhaps also some producers in the United States. In contrast, as we discuss later, a significant realignment of the yuan has the potential not only to reallocate U.S. imports among trading partners but also to affect U.S. macroeconomic conditions and thus the overall size of the trade and current account deficits.

**U.S. trade policy toward apparel and textile imports from China**

U.S. bilateral agreements with China as well as the terms of China’s WTO accession provide for special safeguards aimed at Chinese exports of apparel and textiles. Until its ten-year phaseout was completed on January 1, 2005, the international Multi-Fiber Arrangement (MFA) set a quantitative limit on each country’s exports of each individual textile and apparel product to each importing country. During the phaseout, China made important gains in the U.S. market at the expense of other developing countries, as well as at the expense of U.S. domestic production (see figures 5 and 6); imports from China surged in the early months of 2005. New restrictions on U.S. imports from China thus enjoy the support not only of competing domestic producers, but also of established foreign suppliers in Asia, the Caribbean and Central America, and Africa.
Recent U.S. efforts to limit apparel and textile imports from China are the latest manifestation of a long tradition in U.S. trade policy. Textile imports from Japan had already begun to threaten U.S. producers before World War II, and the United States responded by implementing Japan-specific trade restrictions. In 1956, the United States negotiated a voluntary export restraint on Japanese cotton textile products, resulting in an increase in U.S. imports from other suppliers and of other fibers. International efforts to control this diversion of trade started with the Short-Term and Long-Term Cotton Textile Arrangements (1961–73) and eventually culminated in the Multi-Fiber Arrangement, which regulated most world trade in textile products from 1974 until the end of 1994. These various arrangements meant that international trade in textile products was highly distorted. The system was perceived to be inefficient and unfair. This contrasted sharply with international trade in other manufactured goods, which had been governed since 1947 by rules established under the General Agreement on Tariffs and Trade (GATT).

Lengthy international negotiations completed in 1994 established the World Trade Organization, which replaced the General Agreement on Tariffs and Trade. At the same time, negotiators agreed to bring trade in textile products gradually into conformity with the GATT/WTO system’s basic rules. From 1995 until the end of 2004, trade in textiles and clothing was to be covered by the Agreement on Textiles and Clothing (ATC), which gradually phased out the MFA’s system of quantitative restrictions. When the ATC was being negotiated, the idea of scrapping the MFA enjoyed strong support from developing countries, which included many countries with established or potential comparative advantage in these products; the apparel sector in particular has long been the first step on the road to industrial development. Elimination of all quantitative restrictions on trade in apparel and textiles, thus bringing these products under standard GATT/WTO rules, was expected to provide significant benefits to developing countries. In fact, the agreement to phase out the MFA was widely viewed as a key element in the negotiations, a long-awaited change valuable enough to induce developing countries to accept new rules on services and a more stringent system to protect intellectual property rights.

In the early 1990s, when developing country negotiators “won” ultimate elimination of the MFA, they did not anticipate the rapid pace of China’s integration into the global economy and its acceptance in 2001 as a member of the WTO. As China has emerged as the most important exporter of many apparel and textile products and begun to enjoy the most favored nation (MFN) status afforded to all WTO members, other developing-country exporters have become active, first in urging a delay in full elimination of the MFA and more recently in attempting to maintain established markets in the United States.

Exports from China have made substantial inroads in the U.S. market despite China-specific import barriers. However, the effect of these barriers can be seen by comparing China’s 2003 share in total U.S. imports of textiles and clothing with China’s share in Australia and Japan, countries at a similar stage of development to the United States. These countries offer a useful comparison because, unlike other industrialized countries, Australia and Japan were not using country-specific quotas to restrict textile and apparel imports in 2003. Although both countries applied tariffs to imports of textiles and apparel, import tariffs still allowed lower-cost producers, such as China, to capture larger shares of their import markets. Indeed, as table 1 indicates, in 2003 China accounted for a significantly smaller share of total imports in the United States than in the two other markets. Neither country offers an ideal comparison with the United States because of their greater geographical proximity to China. Still, if Chinese firms’ penetration of the Australian and Japanese import markets is roughly indicative of China’s comparative advantage in textile and clothing production, Chinese exporters have the potential to achieve a substantial further increase in their U.S. market share. As in the past, this gain in market share would likely come at least in part at the expense of other exporters.

Competing exporters can maintain their established shares in the U.S. market by arranging to receive lower tariffs on their exports of textiles and clothing than those applied to exports from China. This can be achieved in at least two different ways. The first is by negotiating a preferential trade agreement with the United States in which each party agrees to give preferential market access to the other, usually along with other concessions. This type of preferential treatment gives some exporters to the U.S. market an advantage in competition with Chinese exporters. As table 2 shows, the United States has recently concluded such negotiations with countries in the Andean region (ATPDEA), Central America (CAFTA–DR), Cambodia, Bahrain, and Morocco. The table also indicates that most of these countries may have been motivated to negotiate such an agreement at least partly because a substantial fraction of their exports to the United States is in textiles and clothing, sectors in which they face increased competition from China for the U.S. market.
China-specific import barriers

A second way other foreign suppliers can face lower tariff rates in the U.S. market than their Chinese competitors is if the U.S. government raises trade barriers against Chinese producers alone. There are several trade policies, all consistent with WTO rules, through which this is possible: antidumping duties, the standard safeguard, the China safeguard, and the China-specific textile and apparel safeguard.\(^\text{20}\) For textiles and apparel, the most frequently used policy has been the China-specific textile and apparel safeguard, which is administered by the Office of Textiles and Apparel (OTEXA) in the U.S. Department of Commerce. Table 3 indicates a number of investigations of Chinese exports of various textile and apparel products undertaken by the OTEXA in 2004. Domestic textile and apparel producers in the United States are not the only potential beneficiaries from imposition of safeguards on imports from China. Because producers in other developing countries may have higher costs than their Chinese counterparts and yet still enjoy a large cost advantage relative to U.S. producers, gains to these other foreign suppliers may be even greater than gains to U.S. producers.

A particular concern regarding the OTEXA process is its lack of transparency. In contrast to the administration of other U.S. trade remedies, including antidumping, the standard safeguard, and the China safeguard, the quasi-judicial and independent U.S. International Trade Commission plays no role in reviewing applications for the China-specific textile and apparel safeguard. Instead, the entire process is carried out internally by the OTEXA in the U.S. Department of Commerce. Because the OTEXA mandate is specifically to assist domestic textile and apparel producers, the OTEXA’s decisions may err on the side of protection by giving little weight to the costs imposed on domestic consumers of the affected products.

The China-specific textile and apparel safeguard provision administered by the OTEXA is due to expire in 2014. However, the U.S. antidumping law can also be used to impose country-specific protection. Traditionally, there has been no need to deal with imports of apparel and textiles through antidumping, since imports were already being managed through the MFA, the WTO Agreement on Textiles and Clothing, and, most recently, the China-specific textile and apparel safeguard. But in other industries, the application of U.S. antidumping duties has had a disproportionate effect on imports from China.\(^\text{21}\) Evidence from table 4 indicates how producers in other industries have managed to use the antidumping provisions to gain an advantage over Chinese exporters through imposition of country-specific protection.

Over the past 15 years, China has been investigated more often and has faced antidumping duties on more products than any other country.\(^\text{22}\)

Antidumping duties are designed to restrict imports of products supplied to the U.S. market at a price below their cost of production or below the price the same firms charge in a foreign market. Most economists believe antidumping duties are usually harmful to overall national well-being. Nonetheless, such duties often find political support because they can be used to limit the competitive pressure in high-cost domestic industries from low-cost foreign competitors. Table 4 lists the average antidumping duty imposed on each of the top ten country targets of U.S. antidumping actions after affirmative investigations. The table indicates that the typical antidumping duty imposed on Chinese firms is much higher than the duty facing firms from other countries. Thus, if an exporter in another country is confronted with a U.S. antidumping measure at the same time as a Chinese exporter, what matters is the size of its duty relative to that levied on a lower-cost supplier in China.\(^\text{23}\)

---

**TABLE 1**

<table>
<thead>
<tr>
<th>Exporting country</th>
<th>Share of U.S. market</th>
<th>Share of Australia’s market</th>
<th>Share of Japan’s market</th>
</tr>
</thead>
<tbody>
<tr>
<td>China</td>
<td>17.4</td>
<td>49.9</td>
<td>71.7</td>
</tr>
<tr>
<td>Mexico</td>
<td>9.9</td>
<td>0.3</td>
<td>0.1</td>
</tr>
<tr>
<td>EU12</td>
<td>6.0</td>
<td>9.5</td>
<td>9.1</td>
</tr>
<tr>
<td>Hong Kong</td>
<td>4.5</td>
<td>2.1</td>
<td>0.3</td>
</tr>
<tr>
<td>India</td>
<td>4.3</td>
<td>3.6</td>
<td>1.1</td>
</tr>
<tr>
<td>Canada</td>
<td>4.1</td>
<td>0.4</td>
<td>0.1</td>
</tr>
<tr>
<td>Korea</td>
<td>3.5</td>
<td>2.7</td>
<td>2.6</td>
</tr>
<tr>
<td>Taiwan</td>
<td>2.9</td>
<td>2.6</td>
<td>1.4</td>
</tr>
<tr>
<td>Pakistan</td>
<td>2.7</td>
<td>2.3</td>
<td>0.3</td>
</tr>
<tr>
<td>New Zealand</td>
<td>0.1</td>
<td>6.9</td>
<td>0.1</td>
</tr>
<tr>
<td>U.S.</td>
<td>44.6</td>
<td>15.9</td>
<td>10.4</td>
</tr>
<tr>
<td>Others</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: Standard International Trade Classification (SITC) codes 26 (textile fibers), 65 (textile yarn, fabrics, and made-up articles), and 84 (articles of apparel and clothing accessories). The European Union 12 (EU12) comprises Belgium, Denmark, France, Germany, Greece, Ireland, Italy, Luxembourg, the Netherlands, Portugal, Spain, and the United Kingdom. Source: Organization for Economic Cooperation and Development.
TABLE 2

<table>
<thead>
<tr>
<th>Country</th>
<th>Textiles and apparel as percent of total exports to U.S.</th>
<th>Textiles and apparel as percent of total imports by U.S.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andean Trade Promotion and Drug Eradication Act (ATPDEA) countries</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Colombia</td>
<td>8.61</td>
<td>0.66</td>
</tr>
<tr>
<td>Bolivia</td>
<td>18.53</td>
<td>0.04</td>
</tr>
<tr>
<td>Peru</td>
<td>21.48</td>
<td>0.62</td>
</tr>
<tr>
<td>Ecuador</td>
<td>0.74</td>
<td>0.02</td>
</tr>
<tr>
<td>Central American-Dominican Republic Free Trade Agreement (CAFTA-DR) countries</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dominican Republic</td>
<td>48.23</td>
<td>2.58</td>
</tr>
<tr>
<td>El Salvador</td>
<td>86.90</td>
<td>2.11</td>
</tr>
<tr>
<td>Honduras</td>
<td>77.78</td>
<td>3.10</td>
</tr>
<tr>
<td>Nicaragua</td>
<td>62.98</td>
<td>0.58</td>
</tr>
<tr>
<td>Guatemala</td>
<td>60.55</td>
<td>2.15</td>
</tr>
<tr>
<td>Costa Rica</td>
<td>17.84</td>
<td>0.72</td>
</tr>
<tr>
<td>Cambodia</td>
<td>99.12</td>
<td>1.51</td>
</tr>
<tr>
<td>Morocco</td>
<td>19.49</td>
<td>0.09</td>
</tr>
<tr>
<td>Bahrain</td>
<td>49.64</td>
<td>0.23</td>
</tr>
<tr>
<td>China</td>
<td>9.80</td>
<td>17.87</td>
</tr>
</tbody>
</table>


The higher average duty for Chinese exporters reflects the methodology that the U.S. Department of Commerce uses to compute the dumping margin, which is the estimated difference between a product’s sale price and its “fair value” (Blonigen, 2003). Commerce Department officials are permitted to choose among alternative methods of calculating the dumping margin, but in the case of China, it is typically the difference between the price charged by Chinese exporters in the U.S. market and an estimate of the Chinese firm’s cost of production. However, because of China’s non-market-economy (NME) status and resulting claims that input prices do not reflect true costs, the Commerce Department frequently uses input prices from “proxy” countries with similar characteristics to estimate the Chinese firms’ costs. The Commerce Department is able to use discretion in the choice of comparison countries, officials can pick countries so as to generate high cost estimates and thus maximize the benefit to the petitioning domestic industry and also to exporters in other countries.

TABLE 3

<table>
<thead>
<tr>
<th>OTEXA category</th>
<th>Product under investigation</th>
</tr>
</thead>
<tbody>
<tr>
<td>349/649</td>
<td>Brassieres and other body supporting garments</td>
</tr>
<tr>
<td>350/650</td>
<td>Dressing gowns and robes</td>
</tr>
<tr>
<td>222</td>
<td>Knit fabric</td>
</tr>
<tr>
<td>447</td>
<td>Wool trousers</td>
</tr>
<tr>
<td>620</td>
<td>Other synthetic filament fabric</td>
</tr>
<tr>
<td>301</td>
<td>Combed cotton yarn</td>
</tr>
<tr>
<td>352/652</td>
<td>Cotton and man-made fiber underwear</td>
</tr>
<tr>
<td>338/339</td>
<td>Men’s &amp; boys’ and women’s &amp; girls’ cotton knit shirts and blouses</td>
</tr>
<tr>
<td>340/640</td>
<td>Men’s &amp; boys’ cotton and man-made fiber shirts, not knit</td>
</tr>
<tr>
<td>638/639</td>
<td>Men’s &amp; boys’ and women’s &amp; girls’ man-made fiber knit shirts and blouses</td>
</tr>
<tr>
<td>647/648</td>
<td>Men’s &amp; boys’ and women’s &amp; girls’ man-made fiber trousers</td>
</tr>
<tr>
<td>347/348</td>
<td>Men’s &amp; boys’ and women’s &amp; girls’ cotton trousers</td>
</tr>
</tbody>
</table>

A high calculated margin may also result when foreign firms fail to respond to the U.S. antidumping investigation process—perhaps because of the short time deadlines, language barriers, or unfamiliarity with the U.S. legal system, all problems that may be particularly acute for a new market entrant like China. In this instance, the Commerce Department is authorized to use the “Facts Available” or the “Best Information Available” in constructing a measure of the foreign firm’s costs—these may be the petitioning domestic industry’s own (high) estimates of those costs.

**Implications of restrictions on U.S. imports from China**

Restrictions on imports from China impose higher costs on U.S. consumers, who now have to pay more for products. However, contrary to the usual justification for these measures, U.S. workers employed in the industry may not be protected. As data in Levinsohn and Petropoulos (2001) illustrate for textiles and apparel, these highly protected industries are characterized by ongoing entry of new firms (and hiring of new workers) at the same time that current plants are closing (and laying off current workers). Protection may increase profitability of the domestic industry, thus encouraging new investment at the same time that older plants, often in a different part of the country, are closing. Although protection can allow the domestic industry to “survive” for a longer period, there is little reason to expect current workers to benefit. The high entry and exit rates reported by Levinsohn and Petropoulos (2001) suggest that protection, by raising the industry’s profitability, may actually accelerate new entry and relocation, thus adding to competitive pressures on current plants and thereby speeding rather than retarding layoffs at these plants. The largest gains from protection are likely to accrue to capital owners who can lower costs by substituting technology and capital for labor in the production process. Abroad, exporters with implicit or explicit preferential access to the U.S. market relative to China will also gain.

Moreover, U.S. imposition of barriers toward imports from China is likely to have substantial international repercussions. China-specific import barriers could cause trade deflection, that is, the redirection of restricted exports to other import markets. There is empirical evidence of trade deflection associated with trade actions limiting U.S. imports from Japan (Bowen and Crowley, 2005a). The import surges in other markets may in turn contribute to new protection in those other markets, as officials use their own antidumping or China-specific safeguard laws to cope with actual or threatened import surges (Messerlin, 2004).

In some cases, however, countries may instead choose to benefit from access to deflected, and thus cheaper, imports as well as an increased variety of available products. To the extent that the deflected imports are intermediate inputs, such as auto parts or steel, the lower prices and greater variety will enhance the competitiveness of foreign consuming firms relative to consuming firms in the United States, which are typically in the higher value-added industries that require these inputs.

With the United States and other trading partners (for example, the European Union and Brazil) imposing or threatening to impose safeguards on Chinese

---

**TABLE 4**

<table>
<thead>
<tr>
<th>Country</th>
<th>Number of antidumping investigations</th>
<th>Number of investigations resulting in duties</th>
<th>Mean duty, conditional on duties imposed, percent</th>
<th>Percent of total U.S. imports</th>
<th>Rank among U.S. imports sources, 1996</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. China</td>
<td>91</td>
<td>61</td>
<td>127.02</td>
<td>3.5</td>
<td>8</td>
</tr>
<tr>
<td>2. Japan</td>
<td>53</td>
<td>33</td>
<td>68.44</td>
<td>14.0</td>
<td>2</td>
</tr>
<tr>
<td>3. Korea</td>
<td>39</td>
<td>20</td>
<td>16.65</td>
<td>2.7</td>
<td>10</td>
</tr>
<tr>
<td>4. Taiwan</td>
<td>30</td>
<td>15</td>
<td>20.46</td>
<td>3.7</td>
<td>7</td>
</tr>
<tr>
<td>5. Mexico</td>
<td>26</td>
<td>11</td>
<td>41.18</td>
<td>10.0</td>
<td>3</td>
</tr>
<tr>
<td>6. Germany</td>
<td>26</td>
<td>10</td>
<td>37.60</td>
<td>4.9</td>
<td>4</td>
</tr>
<tr>
<td>7. India</td>
<td>25</td>
<td>11</td>
<td>52.89</td>
<td>0.8</td>
<td>24</td>
</tr>
<tr>
<td>8. Canada</td>
<td>25</td>
<td>6</td>
<td>25.35</td>
<td>21.0</td>
<td>1</td>
</tr>
<tr>
<td>9. Brazil</td>
<td>24</td>
<td>12</td>
<td>76.47</td>
<td>1.2</td>
<td>16</td>
</tr>
<tr>
<td>10. Italy</td>
<td>19</td>
<td>10</td>
<td>22.75</td>
<td>2.3</td>
<td>11</td>
</tr>
<tr>
<td>Other</td>
<td>272</td>
<td>105</td>
<td>54.55</td>
<td>35.9</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>630</td>
<td>294</td>
<td>64.15</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>

Sources: Data compiled by the authors from the Federal Register; U.S. import data from Feenstra (2000).
apparel imports, in May 2005 the Chinese government responded by proposing export restrictions on the same products, thus “voluntarily” restraining their exports (Barboza and Bradsher, 2005). How would the effects of an export-restricting Chinese policy (for example, an export tax) compare with those of a U.S. import restriction (for example, an import tariff)? Either a Chinese export tax or a U.S. safeguard tariff on imports would reduce trade volume and raise the prices U.S. consumers pay for imported products. But under a Chinese export tax, the Chinese government would collect tax revenue equal to the size of the tax multiplied by the volume of exports. Under a U.S. import tariff, the U.S. government would collect the tax revenue. The standard analysis of a U.S. safeguard tariff on apparel imports predicts that it will reduce the well-being of the economy as a whole, even taking into account possible gains for competing domestic producers and tax revenue generated. If the trade restriction takes the form of a Chinese export tax rather than a U.S. import tariff, the negative impact on U.S. economic well-being would be even larger because the revenue is collected by the Chinese government rather than the U.S. Treasury. For similar reasons, the loss to China is smaller with an export tax than an import tariff.

In the 1970s and 1980s, the United States negotiated voluntary export restraint (VER) agreements with Japan and other highly competitive new exporters as a way to reduce U.S. imports selectively without disrupting relations with established suppliers. As VERs are now prohibited by WTO rules, China’s actions in voluntarily restraining its exports may be seen as inconsistent with China’s obligations under the WTO (Bown, 2002). However, such actions are unlikely to be challenged within the WTO because of the self-enforcing nature of the system. Instead, the export tax revenue may be viewed as implicit compensation paid by the United States to China for imposing a restriction on its own exports.

Yuan appreciation

From June 1995 until July 2005, China pegged the value of its currency at 8.28 yuan per U.S. dollar. In the mid-1990s, this nominal rate approximated an equilibrium market rate (the rate at which the market demand for yuan was equal to the market supply). But recently the demand for yuan at this fixed price greatly exceeded the supply, obliging the Chinese central bank to intervene to meet the excess demand for yuan. Operationally, maintaining the dollar peg required China to buy foreign exchange at the official rate, that is, to supply yuan in exchange for U.S. dollars or other reserve currencies to meet any excess demand for yuan. In the absence of these official transactions, market forces would have caused the yuan’s value to appreciate.

The demand for yuan is derived from foreign demand for Chinese goods and services and also for portfolio and direct investments in China. The supply of yuan is likewise derived from China’s own purchases of foreign goods, services, and assets. Recent upward pressure on the yuan, and the associated surge in China’s official reserves, reflects not only China’s export performance but also strong net inflows of foreign capital, including speculative inflows. In 2004, a record $55 billion of net foreign direct investment plus substantial portfolio capital inflows contributed to the demand for yuan that had to be met by purchases of foreign exchange. By the end of 2004, China’s official foreign-exchange reserves stood at $610 billion, with a substantial share in the form of U.S. dollar assets. Holdings of U.S. Treasury securities accounted for an estimated $194 billion, making China second only to Japan (with $712 billion) in recorded foreign holdings of U.S. Treasuries (Prasad and Wei, 2005). A pegged exchange rate is one of many options available to members under current International Monetary Fund (IMF) rules, and China was hardly alone among developing countries in its choice to peg its currency to the U.S. dollar. It would also be hard to argue that China, which was relatively untouched by the 1997–98 financial crisis that severely affected other countries in the region, could have turned in a better performance under a different currency regime. For two years, Chinese officials rebuffed U.S. efforts to influence China’s currency policy. However, on July 21, 2005, China announced a new policy that responded—at least in qualitative terms—to recent U.S. pressure for yuan appreciation and a more flexible exchange rate regime.

Under the new policy, the yuan will be pegged to a currency basket, another option permitted under current IMF rules and in use elsewhere. The policy change entailed an immediate increase in the dollar value of the yuan by 2.1 percent, and the possibility of further increases relative to the currency basket by as much as 0.3 percent per day. Although a statement from China’s central bank denied any plans for further revaluation in the near future (Barboza and Fuerbringer, 2005), Federal Reserve Chairman Alan Greenspan and U.S. Treasury Secretary John Snow nonetheless welcomed the Chinese decision as a constructive first step—presumably a first step toward a significantly larger appreciation of the yuan relative to the dollar (Andrews, 2005). But China will not make the decision to permit a more significant revaluation lightly. Significant appreciation could undercut export momentum.
and might slow inflows of foreign direct investment as costs of production in China rise relative to those in other potential host nations. Appreciation of the yuan also means a fall in the local-currency value of China’s huge holdings of U.S. dollar assets.\footnote{27}

If the yuan does continue to move upward relative to the U.S. dollar, can the United States expect to see a reduction of its trade deficit? Most answers to this question focus on the resulting change of relative prices. Depending on the extent to which a higher international value of the yuan is passed through to U.S. consumers in the form of higher dollar prices of Chinese goods, substantial yuan appreciation could indeed reduce the U.S. bilateral trade deficit with China. However, this improvement would come mostly through substitution by U.S. importers of goods from other foreign sources, with the pattern depending partly on which nations follow China in allowing their own currencies to appreciate. The effect on the overall U.S. trade deficit could therefore be minor. Yuan appreciation and possible follow-the-leader appreciation by China’s competitors would likewise reduce the local currency prices of some U.S. exports to those markets, thus shifting some international demand toward U.S. producers.

Because the U.S. overall trade deficit must be the difference between U.S. goods production and U.S. expenditure for goods, an overall improvement can occur only to the extent that macroeconomic conditions change so as to raise U.S. production relative to expenditure. Shifting demand toward U.S. suppliers could raise aggregate U.S. production only to the extent that new demand falls on sectors with capacity to expand without significant price increases. Higher import prices could reduce total aggregate expenditure through a wealth effect that makes Americans feel poorer and thus less willing to consume and more willing to produce.

However, a significant appreciation of the yuan relative to the dollar could also affect the U.S. economy through a second and more fundamental channel: by reducing the rate of China’s accumulation of foreign exchange reserves and purchases of U.S. dollar assets. An unanticipated reduction in the demand for dollar-denominated securities would likely put upward pressure on U.S. interest rates. Higher U.S. interest rates would in turn be expected to moderate domestic spending and thereby reduce U.S. demand for all goods and services, not just imports from China. The implication is that, through its effect on relative prices and U.S. capital markets, substantial yuan appreciation could indeed reduce both the bilateral trade deficit and the overall trade deficit.\footnote{28}

**Conclusion**

Is the U.S. trade deficit made in China? The short answer is no. In recent years U.S. bilateral trade deficits have been increasing not only vis-à-vis China but most other trading partners as well (Mann, 2005); the imbalance on bilateral trade with China, though large, still accounts for less than one-quarter of the total. However, official purchases of U.S. dollar assets by China and other U.S. trading partners have facilitated the continued high level of U.S. domestic spending, and some analysts believe that large foreign purchases of U.S. assets have been responsible for keeping U.S. long-term interest rates low (Krugman, 2005; Setser and Roubini, 2005).

Selective trade policies targeting China, such as the recently implemented restrictions on apparel imports, will do little but redistribute U.S. imports across trading partners or products. Moreover, textiles and apparel are likely to be just the beginning of Chinese in­cursions into U.S. markets and perhaps also just the beginning of U.S. trade policy actions toward China. China already has substantial export growth in steel and auto parts, thus entering into competition with two additional domestic industries that have benefited historically from import protection. If the United States also moves to protect these industries from Chinese competitors, how might other exporters to the U.S. market respond? In contrast to the case of apparel and textiles, in the steel and auto sectors China’s main competitors in the U.S. market are not just other developing countries but also the European Union, Japan, Korea, and other industrialized countries. Will these countries attempt to negotiate preferential trade agreements with the United States? Will a system evolve along the lines of the 2002 steel safeguard (Bown, 2004), in which U.S. tariffs were applied on a non-preferential (MFN) basis, but with substantial product exclusions favoring specific foreign suppliers? Will affected firms increase foreign direct investment in the United States, thus allowing their U.S. subsidiaries to use the domestic protection-seeking process from an insider’s perspective?

Like China-specific trade policies at the level of individual products, a stronger yuan may benefit other U.S. trading partners, especially ones that have recently lost market share due to depreciation of the U.S. dollar against their own currencies. But yuan appreciation also means lower Chinese official purchases of U.S. dollars, which would likely put upward pressure on U.S. interest rates and thus moderate domestic spending. So yuan appreciation could reduce the U.S. overall trade imbalance, although it is not necessary to address the fundamental cause of the deficit, which lies in current macroeconomic conditions at home.
Once a country’s trade balance is in deficit or surplus, even equal growth rates of exports and imports would mean an equal rate of growth of the trade imbalance. However, the ratio of the trade balance to gross domestic product (GDP) would remain constant if output also grew at the same rate. To reduce the size of a trade deficit, exports would need to grow at a faster rate than imports; to reduce the trade deficit to GDP ratio, it would be sufficient for exports and imports to grow at the same rate as long as output was growing faster.

The balance on goods and services trade adds exports and imports of services (for example, international shipping services and the services of engineering or consulting firms) to the merchandise trade balance. The current account balance adds two further items: net foreign income (mostly interest and dividends from U.S. assets abroad and foreign assets in the United States) and unilateral transfers (including remittances, private charitable contributions, and foreign aid). Thus, the current account deficit summarizes the net of all non-asset spending and receipts over a given period—the final bill the nation must settle.

Although the merchandise trade balance is a relatively narrow measure of U.S. trade performance, we focus on it here because, empirically, it is the largest component of the current account (see figure 3, p. 5), and proposals motivated by the size of the overall current account deficit often seek to use trade policy to reduce the merchandise trade deficit.

This figure is based on U.S. government data. Chinese customs data give a much smaller figure. The discrepancy is due to differences in the statistical treatment of trade that passes through Hong Kong. Adjustments by Feenstra et al. (1999) and Schindler and Beckett (2005) yield estimates of the bilateral deficit that lie between the official U.S. and Chinese figures.

On July 21, 2005, the Chinese government announced that it would no longer peg the yuan to the U.S. dollar. The new policy entailed an immediate appreciation of 2.1 percent of the yuan relative to the dollar and the possibility of further increases over time.

Total domestic goods expenditure can be broken down into private consumption and investment spending plus purchases by all levels of government.

Empirical research has shown that trade barriers targeting specific countries (Moore, 1996; Prusa, 2001; Bown, 2004) or specific products (Feenstra, 1988; Goldberg, 1995) cause similar imports from other countries or in other (higher quality) product classes to rise; Greenspan (2005) emphasized this substitution in remarks before the U.S. Senate Finance Committee. Because induced imports offset the decline in targeted imports, the restrictions produce little or no effect on the overall level of imports or the size of the trade deficit. A number of theoretical papers (Razin and Svensson, 1983; van Winbergen, 1987; Gardner and Kimbrough, 1989) derive conditions under which tariffs that are broadly applied to imports from all countries and in all sectors can reduce the overall trade deficit. These authors emphasize that even broadly applied tariffs can reduce the trade deficit only in specific circumstances.

This interpretation is loose because not all transactions recorded in the financial account represent lending and borrowing. Financial account transactions include changes in equity positions. These differ from loans in that there is no predetermined schedule of interest or principal payments. For example, the construction of a U.S. factory or the purchase of a U.S. department store by a foreign firm would appear as capital inflows along with loans to U.S. borrowers or purchases of U.S. bonds by foreign banks. Moreover, a country with a financial account surplus may be selling off foreign assets accumulated in previous years rather than borrowing abroad.

Opinions may differ as to when this point is reached and what to do about it. The U.S. Trade Deficit Review Commission failed to reach consensus on the causes and consequences of the U.S. trade and current account deficits. Most chapters of the Commission’s final report included separate sections detailing the conflicting views of the Democratic and Republican commissioners (U.S. Trade Deficit Review Commission, 2000).

As Collins (1999) emphasizes, the trade balance is determined together with other macroeconomic variables, such as output growth and job creation.

Although most discussion of sustainability has focused on the current account, some authors also derive sustainable ratios of the deficit on trade in goods and services to GDP, which is currently about 4 percent. Kourapisas (2003) calculates a maximum sustainable ratio of 1.4 percent; Robbini and Setser (2004) derive a maximum ratio of 1 percent. Because the United States is a net exporter of services, these figures would imply a slightly higher sustainability ratio for the merchandise trade deficit to GDP.

Japan’s bilateral trade surplus with the United States is only half as large as China’s, while Japan’s economy is more than twice as large as China’s. On U.S. bilateral trade with other Asian countries, see Hufbauer and Wong (2004).

A survey article in The Economist reports estimates of yuan undervaluation ranging from over 40 percent to as little as 7 percent (Economist Newspaper Limited, 2005). The small appreciation on July 21, 2005, leaves the yuan undervalued even in terms of the lower bound of these estimates.

A related concern is that many Chinese businesses do not respect intellectual property rights. U.S. business groups claim that ongoing violations—patent and copyright infringement as well as outright counterfeiting and piracy—are costing them billions of dollars each year in foregone sales or royalty revenue (Office of the U.S. Trade Representative, 2005). Although China is similar to many other developing countries in its failure to control unauthorized use of intellectual property, the size of the Chinese market means that estimated losses are much larger.

Hoekman and Kostecki (1995) provide a detailed history of negotiations preceding the establishment of the WTO.

Most favored nation treatment, that is, nondiscrimination among trading partners in terms of the barriers to their exports, is a central element of WTO rules. In practice, however, MFN treatment is subject to several important exceptions. See Crowley (2003).

Australia abolished textile and apparel quotas in favor of import tariffs in 1991 (Garnaut, 2002). Japan never applied country-specific quotas to imports of textiles and apparel.
Although WTO rules prohibit most tariff increases that target exports from a specific country, the rules permit tariffs used to implement antidumping and safeguards.

For details of China-specific U.S. trade measures and their effects, see Bown and McCulloch (2005).

Data from Zanardi (2004) indicate that this result is not specific to the United States; Chinese firms have been targeted most frequently by all users of antidumping policies worldwide.

This assumes that the U.S. antidumping duty faced by non-Chinese exporters is not high enough to redirect all import demand to domestic producers.

Exporters may obtain preferential access through a preferential trade agreement, or if lower antidumping duties are levied on their products, or if a China-specific safeguard is imposed.

This means the nominal exchange rate remained fixed. In real terms, the yuan depreciated about 2.4 percent relative to the dollar over this period, because cumulative inflation in China was slightly higher than in the United States (Hanke and Connolly, 2005).

As of August 2005, the Chinese government had indicated the composition of the basket but not the weights or other details of the operation of the new currency regime. In practice, this means China has broad latitude in determining future movements of the yuan.

Yuan appreciation could put further pressure on the already weak Chinese banking system to the extent that those banks hold U.S. dollar-denominated assets or have made loans for which repayment depends on export revenues.

Upward pressure on U.S. interest rates due to reduced Chinese purchases of U.S. Treasury bonds would tend to raise demand from other investors, thus moderating the effect on U.S. rates.

REFERENCES


Bown, Chad P., 2004, “How different are safeguards from antidumping? Evidence from U.S. trade policies toward steel,” Brandeis University, manuscript, July.


Economist Newspaper Limited, 2005, “Precisely wrong: China’s currency may not be as cheap as is commonly believed,” Economist, June 25, p. 76.


Setser, Brad, and Nouriel Roubini, 2005, “How scary is the deficit? Our money, our debt, our problem,” Foreign Affairs, July/August.


Energy markets and the Midwest economy

Richard H. Mattoon

Introduction and summary
The price and availability of energy has long been a critical concern to industrialized nations. In 2002, the industrialized nations as represented by the Organization for Economic Cooperation and Development (OECD) were responsible for consuming 61 percent of the world’s petroleum and 51 percent of the world’s natural gas. Higher prices have lead to concerns about the potential drag that energy costs might have on economic growth since the first oil embargo of the 1970s. For energy-intensive industries, high energy prices can be particularly destructive. Yet the recent disruptive impact of high energy prices appears to be muted as the U.S. economy has shifted to a service base and gains have been made in reducing the reliance on energy to produce U.S. output. Twenty-five years ago it took 15,000 Btu (British thermal units) to produce $1 worth of the U.S. gross domestic product (GDP). By 2003, this had fallen to 9,500 Btu, a decline of nearly 37 percent.

For much of 2004–05, a significant economic story was the rising price of virtually all types of energy. While crude oil prices grabbed the headlines as nominal spot prices hit record highs approaching $60 a barrel, less attention has been paid to the rise in natural gas and coal prices. The spot price for natural gas has risen from $4/MMBtu (per thousand Btu) in November 2003 to better than $6.25 by early November 2004. In late October 2004, spot prices peaked at nearly $8. Coal prices have seen an even more dramatic run up in the last several years. The spot price of Central Appalachian coal has risen from roughly $28 per ton in the first part of 2002 to a record nominal price of $66.50 a ton by late October 2004.

More than the rest of the U.S., the Midwest, with its industrial legacy and seasonal weather pattern of cold winters and hot summers, is more energy reliant than the rest of the nation. In this article I will examine in greater detail how the Midwest economy (Seventh Federal Reserve District—7G) is exposed to energy prices and how this exposure has changed over time. In particular, I will look at two sources of change. First I will examine systematic improvements in energy efficiency and/or conservation. Second, I will examine changes in the structure of the economy away from energy-intensive industries toward services industries, focusing on how economic structure differs from state to state within the District and in comparison to the U.S. as a whole. I find that the Midwest has followed a similar path as the rest of the nation in reducing the amount of energy input needed to produce $1 of GDP. However, I also find that the region has, on a relative basis, increased its national share of seven energy-intensive industries, suggesting that the region will feel the effects of rising energy prices slightly more than the nation as a whole. In particular, the region is highly reliant on natural gas, so volatility and price increases in this fuel bear our particular attention.

The remainder of this article is organized as follows. The second section will provide a brief literature review regarding the importance of energy markets and energy prices to the macroeconomy and to state and regional economies. The third section will briefly describe the recent evolution of energy prices and volatility that shows that volatility may be the more difficult issue for economic performance. Energy markets have historically had a boom and bust cycle that has discouraged both consumers and producers from changing their behavior. The fourth section describes the economies of the 7-G states and how their similarities and/or differences from other parts of the economy.
country with respect to energy prices. The evidence suggests that the region is still relatively more exposed to energy costs than the rest of the nation as a whole. The final section offers some concluding observations.

**Related literature**

Since the mid-1970s economists have been examining the effect of energy and, in particular, oil price shocks on the macroeconomy. Early empirical studies tended to measure the effect by regressing GDP on oil prices and other selected variables (Rasche and Tatom, 1977a, 1977b). However isolating the effect of oil prices has always been a difficult econometric task providing very little in the way of “clean experiments” where oil prices alone cause declines in economic output. Darby (1982) makes the case that the 1973 oil price shock was the closest to a clean experiment in as much as the world was just emerging from the international monetary arrangements established at Bretton Woods and the U.S. was emerging from a period of generalized price controls, reducing the confounding effect of other factors on economic performance. In a landmark article, Hamilton (1983) found that an oil price increase has preceded every recession in the U.S. since World War II with the exception of 1960. This finding focused research attention on the importance of oil in the economy throughout the post-World War II era rather than just on the economic effect from the two oil embargoes of the 1970s.

Other work examined the effect of price volatility on adjustment mechanism in the economy. Gilbert and Mork (1986) and Mork (1989) were interested in explaining the weaker-oil-price–GDP relationship that Hamilton had found during the late 1970s by examining whether the economy had somehow adjusted to mitigate the impact of oil price shocks. These studies examined whether price movements have symmetric effects on production possibilities. They found that any change in direction triggers resource reallocation; however, increases in prices tend to have a more significant effect on the economy than decreases.

More recently, oil and energy price research has focused on its relationship to business cycle theory. In particular research has focused on the relationship between monetary policy adjustment and oil price changes (Bernake, Gertler, and Watson, 1997). This research builds on the work of Tobin (1980) that questioned whether a resource that accounts for such a small share of U.S. GDP (oil is roughly 3 percent) could cause large losses in GDP in ensuing recessions. These authors suggest that it has been monetary policy adjustments in response to an oil price spike that may have played a larger role in triggering economic decline than the oil price spike in the first place.

A recent review article (Jones, Leiby, and Paik, 2004) suggested some interesting conclusions about what is known about the current relationship between oil prices and GDP. Their review of the literature found that when price movements have been large compared with recent volatility, the effects of oil prices on the economy have been greatest. Sharp volatility is more important because a sustained higher price causes consumers and producers to alter their behavior in response to the higher price. Further they suggest that the effect is mostly seen in the reallocation of labor within specific industries. Reallocation of labor is particularly intense for manufacturing. Davis and Haltiwanger (2001) found that the oil price shock of 1973 was related to a job reallocation of 11 percent of total manufacturing employment over the next 15 quarters. However, they note in particular that this reallocation occurs within industry classifications and even at the plant level. Sector-specific and plant-specific factors are at work, suggesting that the real distributional effect of an energy price shock needs to be examined at the sector or plant level.

**Literature on regional adjustment to energy prices**

Brown and Yucel (2004) have documented the effect of higher oil prices on the Texas economy noting that the region’s industries have become less energy reliant tending to mute the impact of sudden increases in oil prices. However, they have also noted that, as the oil producing and refining industries have declined in importance, the Texas economy has not received the same boost as in the past when oil companies were major beneficiaries of higher prices. Using a vector autoregression (VAR) model, they found that rising oil prices only raised Texas gross state product (GSP) by one-fifth as much during the period 1988 to 2002 as they had during the period 1970–87.

In another regional study, Bradbury (2005) looked at the effect of higher energy prices on households by census region over the winter of 2003–04 in comparison to the winter of 2004–05. During this period the U.S. Department of Energy forecasted that fuel oil prices would increase by 39 percent, gasoline by 24 percent, and natural gas by 13 percent. This study recognized that the relative fuel mix used by a given region largely determines what the effect of increased energy prices will be on the average household. From the household’s perspective, energy consumption falls into two broad baskets—home heating and residential needs and transportation. Each region of the country has specific energy needs determined by weather, driving patterns, and region-specific preferences for certain fuel types. Bradbury finds that the short-run impact of higher energy prices will be most felt in New England.
due largely to a preference for heating oil. For the Midwest, reliance on natural gas and higher than national average transportation needs drives energy costs. The projected energy cost increase as a share of consumer spending from the winter of 2003–04 to the winter of 2004–05 is estimated at 1.26 percent for New England and 1.11 percent for the East North Central region.

**A brief history of fuel prices and volatility**

In this section I provide an overview of fuel price behavior and measures of price volatility. As figure 1 demonstrates, the price movements of the three major energy fuels used in the U.S. have not moved in synch. This should not be of any real surprise given that each fuel is governed by its own set of market dynamics. For example, oil prices reflect supply and demand conditions in a world market. Oil prices were clearly affected by the disruption in supply from the Arab oil boycotts of the 1970s and more recently by the growing world demand (particularly from China and India) and concerns over potential supply disruptions. In contrast, natural gas and coal prices reflect regional conditions and certain idiosyncrasies specific to each fuel. In the case of natural gas, the U.S. market is highly integrated with Canada, creating a regional North American market. Natural gas prices reflect the infrastructure used to deliver the product, and prices are set at regional trading hubs. The recent increase in natural gas prices reflects limitations in the pipeline infrastructure to deliver the product and growing concern that North American gas fields are maturing leading to more expensive extraction and lower well productivity. In addition the relative inability to increase liquefied natural gas (LNG) imports by the U.S. (LNG accounts for 2 percent of U.S. energy consumption) means that no ready substitute for North American production exists. Finally, in the case of coal, the market has been shaped by continued concern over the environmental attributes of the fuel. While coal is still the preferred fuel for baseload electricity generation, years of environmental regulation and potential concern over options such as carbon taxes has limited coal consumption. The recent increase in coal prices reflects a renewed desire by utilities to burn coal to offset the sharp increases in costs of alternative fuels, particularly natural gas.

Figure 2 provides a slightly different look at fuel prices. In it, the prices for inflation have been adjusted and normalized by relating the price to the price per million Btu. Then, to more closely reflect U.S. fuel consumption patterns, motor fuel is substituted

![Figure 1: Average annual energy prices](image1)


![Figure 2: Real fuel prices, dollars per million Btu, 1970–2004](image2)

for oil since nearly 70 percent of U.S. oil consumption comes in the form of motor fuel. Perhaps most notable in this figure is the relatively long period of time in which all three fuels experienced flat or declining real prices—from the mid 1980s to the late 1990s. Not until 2000 do you begin to see a sharp upturn in both motor fuel and natural gas prices, a harbinger of things to come.

Measuring volatility

It is not just price that has a major impact on energy markets. As was noted in the literature review, economists have found that the relative volatility of fuel prices has a significant effect on the response of the economy, households, and firms to sudden changes in energy prices. Casual evidence would suggest that greater relative volatility slows the process of adaptation since neither consumers nor producers know whether to make fundamental behavioral changes in the face of uncertain prices. For consumers, if the price increase is seen as temporary, they are likely to maintain their energy consumption habits by reducing expenditures on other items or reducing savings rather than making significant changes such as installing fuel-efficient appliances or buying a more fuel-efficient car. Producers fearing a boom and bust cycle in energy prices are likely to be wary of making investments in long-lived physical assets based on prices that may be short-lived. Even today in the era of oil at well above $60 a barrel, major oil companies are determining their investment decisions based on a long-run price of oil in the high $20 a barrel range. Figure 3 shows the volatility of the three major fuels measured by the annual standard deviation. Of the three fuels, natural gas has exhibited the largest volatility for more than 30 years.

Sources of volatility

For all fuels, the recent increase in volatility is most closely related to increased world demand and shrinking surplus capacity. In general, the fuel system is operating at a higher capacity, and this can make supplies tight when demand increases since there is little surplus capacity. In the U.S., oil refineries have been operating above 90 percent capacity utilization since the early 1990s leaving little room to compensate for an unplanned shutdown of a refinery. In addition, the increased

![FIGURE 3](chart, at www.eia.doe.gov/emeu/states/_states.html, for fuel prices.)
requirements for reformulated gasoline, now 30 percent of the U.S. motor fuel market, further reduces the flexibility of refineries by requiring the production of specialty motor fuels to meet environmental standards for specific parts of the country. In the natural gas field, it can take up to a year for significant new gas production to come online even in the face of higher prices. Constraints in pipeline capacity also can limit the ability to get gas to the market even if it is available. In the case of LNG, it can take up to ten years to site and build a terminal due to siting restrictions and construction expense. Finally, there is the general reluctance to bring new energy resources online given the long time frame over which the investment must pay out. Energy assets often have useful lives of 20 years to 30 years. The decision to invest in a production asset is determined by the cash flow expected from the asset based on the estimated price of the fuel over the life of the facility. This tends to make energy companies somewhat conservative even when prices are high. Having seen prior booms and busts in prices, these companies’ conservatism is understandable. Energy industry analysts believe that market volatility slows investment by oil and gas companies. The bottom line is that with lower reserves, tighter production, and an inability to rapidly respond to increased demand, price becomes the mechanism for balancing the market in the short run.

Turning attention to Midwest energy and the changing structure of the 7-G economy

Much of the concern over higher energy prices in the Midwest has to do with the region’s economic structure. The region has long been known as the nation’s manufacturing belt. Manufacturing is significantly more energy intensive, so it bears to reason that higher energy costs will disproportionately affect the region’s economy.

Figure 4, panel A–F illustrates the changing structure of the 7-G economy based on the composition of GSP for the individual states and for the District as a whole. What is most striking is that while the share of GSP derived from manufacturing in the District has declined significantly since 1980, it is still well above the U.S. average. In contrast, the District is slightly below the U.S. average in its share of GSP from less energy-intensive industry sectors such as services and finance, insurance, and real estate. Perhaps more interesting is the contrasting structure of each state’s economy. Illinois has dramatically reduced GSP from manufacturing from 25 percent in 1980 to only 13 percent by 2003. The Illinois economy has departed from many of its industrial neighbors and now has a structure that essentially mirrors that of the U.S. In large measure this can be attributed to the restructuring of the Chicago metropolitan economy where manufacturing has declined dramatically and been replaced by growth in business services, retail trade, and convention and tourism. In contrast, Indiana and Wisconsin continue to have significantly higher shares of GSP from manufacturing. In both states, manufacturing is still the largest share of GSP at 27 percent and 22 percent, respectively, in 2003. Over this period, these states have seen less systemic restructuring by industry sector as measured by output. Indiana in particular has maintained a heavy concentration of durable manufacturing in sectors such as recreational vehicles and automotive parts. Iowa and Michigan fall somewhere in the middle. Both have had significant declines in GSP attributed to manufacturing (Iowa fell from 26 percent to 20 percent and Michigan from 31 percent to 21 percent) but they still have manufacturing shares well above the national average. Michigan manufacturing is still highly related to the auto sector.

Another factor increasing the energy dependence of the region is climate. Being a region characterized by cold winters and hot summers, energy demand for heating and cooling in the Midwest is relatively high. One of the easiest ways to document the relatively harsher climate of the 7-G states is through the use of heating and cooling degree days. Heating degree days calculate the daily variation in temperature at a location below 65 degrees Fahrenheit, while cooling degree days calculate the variation above 65 degrees. States with high heating degree totals require significant energy for space heating and usually are marked by high consumption of natural gas and fuel oil. States with high cooling degree totals are usually large consumers of electricity needed to run air conditioners. Table 1 (p. 26) shows the average annual heating and cooling degree totals from 1971–2000 weighted by each state’s population in the 7-G, and for the U.S. population as a whole.

The significant variation in heating days above the U.S. average places a special emphasis on the use of natural gas in the region. As table 2 (p. 26) demonstrates, natural gas is overwhelmingly the preferred heating fuel in the District states, and the region’s cold winters make the Midwest more reliant on natural gas than any other region.

When it comes to energy consumption the five states that compose the Seventh District have differing patterns that tend to reflect the underlying structure of their economies. Table 2 compares energy utilization in each state compared with the U.S.
FIGURE 4

Percent of total gross state product by industry and comparison to the U.S.

A. Seventh District states

B. Illinois

C. Indiana

D. Iowa

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

percent

Percent of total gross state product by industry and comparison to the U.S.
Total energy consumption is above the national average for all states except Iowa and much of this has been attributed to the above national average concentration of energy-intensive manufacturing industries and midwestern climate. However, on a per capita consumption basis, the region appears more moderate in its consumption patterns with the exception of Indiana.

The changing role of energy related to economic output

From an economic perspective, an important trend has been the declining role of energy as an input to producing gross product in the U.S. This trend has been mirrored in the 7-G, as well as played a significant role in reducing the importance of energy as a basic input to production. Figure 5 displays the change in the number of Btus needed to produce $1 of gross product. For all three fuel types, Btu equivalents are used to allow for more accurate comparisons. The declines have been dramatic with the amount of energy needed to produce $1 of gross product dropping by 77 percent for natural gas, 76 percent for motor fuel, and 67 percent for coal. In the case of natural gas, 7-G states followed this pattern for the most part although Indiana, Michigan, and Iowa required higher levels of usage on natural gas to produce GSP. As for coal, it is worth noting the significantly higher utilization of coal to produce GSP. Indiana uses coal as a primary fuel for 80 percent of its electricity generation and is more dependent on coal as a fuel than the rest of the region.

Jones, Leiby, and Paik (2004) found that the largest economic effect of energy price spikes was demonstrated through changes in employment in specific industries. They suggest that reallocation related to an energy price shock is often determined at the plant level making estimates of economic impact at even the broad industry level potentially misleading. In order to test this idea, I have selected
the seven industries (aluminum, chemicals, forest products, glass, metal casting, petroleum, and steel) identified by the U.S. Energy Information Agency as the most energy intensive and examine how employment has changed in these industries following oil shocks. I will do this for the five District states and for the U.S. as a whole. In addition, I will look at the effect of the relative concentration of these industries on the District states over time. Using location quotients (LQ) based on employment shares, I will demonstrate which states in the District have the largest concentration of these energy-intensive industries and how this has changed over time. This will shed light on the question of whether the employment in region has in a relative sense become more or less exposed to energy dependent industries over time.

In evaluating the structure of the Seventh District economy, there is clearly a lack of many energy-producing industries, with the exception of coal; however there is a reasonable concentration of employment in energy-intensive industries. Table 3 (p. 28) shows some basic properties of these industries and their relative concentration in the Seventh District.

As this table demonstrates, Indiana in particular has a concentration of energy-intensive industries. In total, Indiana had nearly $37 billion in shipments from these industries, with total employment of 83,000. In all, these industries made up more than 7 percent of GSP. Individual industries played important roles in specific states. The forest products industry in Wisconsin is responsible for almost 4 percent of that state’s GSP and employs 71,000. In Illinois, chemicals account for almost 2 percent of GSP and employ 58,000. However to assess the impact that high energy prices might have on these industries, it is important to examine what their long-term growth trends have been.

**Employment trend**

Figure 6 (p. 28) shows total employment in these seven industries over 3 decades. In the case of the 7-G states employment decline was more pronounced than the U.S. from 1972 to 1982, however employment turned around in the early 1980s and these industries showed job gains up until 2000. During this period, District employment outperformed the U.S. as a whole. This pattern is more clearly reflected in figure 7 (p. 29) showing the annual percentage change in employment. It is also worth noting the behavior of employment in light of major oil shocks. The first and second oil embargoes of the 1970s and the related price shocks created significant job loss in these seven industries, more so in the District than for the nation as a whole. Interestingly, the change in employment is significantly less volatile following the 1990 Persian Gulf crisis. Some analysts suggest that this reflects a reallocation
of labor in these industries where production has moved off shore. However it is worth noting that the sharper employment decline of these industries in the District states beginning in 2000 may reflect that these industries are still relatively more concentrated in the District than they are in the nation as a whole, and therefore are more likely to respond to higher energy prices.

Evaluating the relative concentration of energy-intensive industries in the 7-G

I have decided to use location quotients (LQ) to examine how the relative concentration of these seven industries has changed in the District based on employment. An LQ is a common measure in economic geography that identifies the relative significance of a phenomenon (in this case employment in energy-intensive industries) in a region or state compared with a benchmark region (in this case the U.S.). In interpreting the results from LQs, any number above 1 indicates an employment share in that industry that is above the national average. For example, the forest products industry in Wisconsin has increased its importance to that state’s economy. Forest products are represented by two Standard Industrial Classifications (SIC), 2400 and 2600. In Wisconsin the location quotient for employment in SIC 2400 rose from 1.16 in 1972 to 1.96 in 2002. This means that Wisconsin has nearly double the national average of employment in this sector. Likewise, employment rose from 18,500 in 1972 to 32,100 in 2002. SIC 2600 experienced even larger growth with its LQ rising from 2.99 to 3.82. Employment rose from 43,000 to 50,000.

Chemicals in Indiana have also grown in importance to that state’s economy. In 1972, chemicals had employment of 26,000 and an LQ of 1. By 2002, employment had risen to 33,000 and the LQ to 1.46.

On an industry-specific basis, all of the industries have seen gains in their LQ since 1972 in the 7-G. Specifically,

- Metal casting, 2.0 in 1972, 2.5 in 2002,
- Steel, 1.6 in 1972, 2.1 in 2002,
- Aluminum, 1.7 in 1972, 2.0 in 2002,
- Glass, 0.9 in 1972, 1.11 in 2002,
TABLE 3

Energy consumption patterns in 7-G and U.S.

<table>
<thead>
<tr>
<th>Industry</th>
<th>Energy intensity (energy purchased as a % of product shipped, 2001)</th>
<th>Primary energy/fuel type used in production</th>
<th>Industry percentage of GSP in top 7-G states</th>
<th>Employment in thousands in top 7-G states (national rank)</th>
<th>Value of shipments in billions of $, 2000 (national rank)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aluminum</td>
<td>6.9</td>
<td>Electricity 76% Natural gas 20%</td>
<td>Indiana .43%</td>
<td>Indiana 6.1 (1)</td>
<td>Indiana $2.5 (2)</td>
</tr>
<tr>
<td>Chemicals</td>
<td>3.7</td>
<td>Natural gas 37%</td>
<td>Indiana 4.6%</td>
<td>Illinois 58.1 (5)</td>
<td>Illinois $22.2 (8)</td>
</tr>
<tr>
<td>Forest products</td>
<td>Wood-4.7 Paper-2.0</td>
<td>Wood residues 50%</td>
<td>Wisconsin 3.9%</td>
<td>Wisconsin 71.3 (2)</td>
<td>Wisconsin $18.1 (1)</td>
</tr>
<tr>
<td>Glass</td>
<td>6.5</td>
<td>Natural gas 54%</td>
<td>Indiana .05%</td>
<td>Indiana 7.2 (9)</td>
<td>Michigan $1.2 (9)</td>
</tr>
<tr>
<td>Metal casting</td>
<td>4.7</td>
<td>Natural gas 37%</td>
<td>Wisconsin .72%</td>
<td>Michigan 7.2 (9)</td>
<td>Michigan 1.6 (8)</td>
</tr>
<tr>
<td>Petroleum</td>
<td>3.9</td>
<td>Refined products—refinery gas, coal, coke, and other 94%</td>
<td>Illinois .4%</td>
<td>Illinois 5.6 (5)</td>
<td>Illinois $14.6 (4)</td>
</tr>
<tr>
<td>Steel</td>
<td>7.7</td>
<td>Natural gas 42% Coal 31%</td>
<td>Indiana 1.75%</td>
<td>Indiana 32.0 (2)</td>
<td>Indiana $12 (2)</td>
</tr>
</tbody>
</table>

- Chemicals, 0.7 in 1972, 1.3 in 2002,
- Forest products 1.25 in 1972 to 1.4 in 2002, and
- Petroleum 0.6 in 1972, 0.7 in 2002.

Only petroleum has a relative employment concentration below the U.S. average and three industries have concentrations that are double the U.S. average.

For the District as a whole, while the relative concentration of energy-intensive industries has increased over this period total employment has declined. In 1972 the LQ for all seven industries district-wide was 1.11 and by 2002 it was 1.24. Employment however fell from 998,600 to 731,000 (26.9 percent). On a national level employment in these seven industries fell from 5.580 million to 4.036 million over the same period—a decline of 27.7 percent.

There is also significant variation by state (see figure 8). For the entire Seventh District, employment in these highly energy-intensive industries was a little better than 13 percent above the nation. However, the LQ for the District has been on the rise since hitting its trough in 1981 and has been consistently above the nation since 1988. On an individual state basis, the story is quite different. Illinois and Michigan have consistently lowered their employment LQs in these industries. Iowa while still
below the national average for employment at 98, has shown rapid gains with its LQ doubling since 1975. The two states that have the largest concentration of employment in these energy sensitive industries are Indiana and Wisconsin. Indiana’s LQ was double the U.S. by 2000 while Wisconsin had seen its LQ rise to 129 from a low of 70 in 1982.

**Conclusion**

This article makes three basic observations about energy markets trends and behavior. First, the market dynamics for individual fuel types are quite different. While oil prices are largely set in a world market, natural gas and coal are influenced by regional dynamics. Issues of fuel security, infrastructure for delivering the fuel, government regulation, and the development of spot markets and trading centers all have varying influences on the behavior of each fuel. However, recently energy prices appear to have become more closely linked. Demand for all fuel types has been on the rise, and fuel substitution has been limited leading to similar levels of increases in all fuels. Second, price volatility appears to influence investment decisions and may discourage investment in costly energy infrastructure. Finally, the Midwest’s economic composition suggests that certain industries (metal casting, steel, and aluminum) and states (Indiana and Wisconsin) will be more significantly impacted by higher fuel prices.

Many extensions to this line of research are possible. Ultimately to properly assess the impact of energy costs or energy spikes on the region’s economy it is necessary to identify the relative importance of energy as a cost of business to individual firms. An old maxim in economics is that high energy prices act like a tax on consumers. If this is true the interesting questions need to focus on the incidence (or distribution) of that tax based on specific attributes of consumers/industries. Further, it must be recognized that the price paid for fuel and energy varies depending on company-specific purchasing agreements. Some companies buy fuel on long-term contracts and some at spot market prices. The impact of reported higher spot market prices may be negligible on a well-hedged fund. In addition more research needs to be done to examine the affect of energy prices on secondary markets for industries. For example, the Midwest is still home to the Big Three auto producers. Reports from Detroit blame high gas prices for reducing demand for large sport utility vehicles. How will this affect the region’s economy? Finally, more needs to be done to understand the impact of energy as it applies to the reallocation of resources globally. Manufacturers increasingly see
their competition arising off shore. Is the energy picture different for industries located in key competitor nations? By answering these questions, we can ultimately develop a far clearer understanding of the impact of energy prices on regional economies.

NOTES
5 U.S. Energy Information Agency (2004). Percent increase reflects the change in the price per ton for central Appalachian coal from May 31, 2002 to October 29, 2004. The October 29 price was a record high of $66.50 per ton.
6 Specifically the study found that a 10 percent increase in oil prices increased Texas gross state product 2.6 percent during the period 1970–87 while only increasing gross state product by 0.4 percent in the period 1988–2002. In addition the authors found that a 10 percent increase in oil prices in the first period increased Texas employment by 1 percent. In the 1988–2002 period a ten percent rise in oil prices lead to a 0.4 percent decline in employment.
7 These estimates are based on forecasted prices in the U.S. Energy Information Agency (EIA) Short-term Energy Outlook for November 2004. The EIA estimated that from the winter of 2004 to the winter of 2005 that the price of no. 2 heating oil would rise by 38.7 percent, residual natural gas by 12.6 percent, propane by 22.4 percent, residential electricity 1.6 percent, and gasoline by 23.9 percent. For the U.S. as a whole, energy is 7.1 percent of the CPI-U. In the case of the Midwest, energy is a slightly higher share of the consumer market basket at 7.4 percent. Residential fuel is responsible for 4.1 percent of the energy cost and motor fuel for the remaining 3.3 percent. Within the residential fuel category, 52 percent of the estimates for the 2003 fuel mix was represented by electricity, 43.3 percent by natural gas, 3.5 percent by propane, and the remaining small shares from fuel oil and kerosene. The Midwest reliance on natural gas in its fuel mix is the highest in the nation. The next closest region is the Mid-Atlantic at 34.1 percent. Clearly changes in natural gas prices will have a larger influence on household budgets in the Midwest. As Bradbury points out, in the short run, there is little evidence that the household sector reacts to higher fuel prices by dramatically reducing consumption or switching to less expensive fuels. This can occur in the long run if higher energy prices appeared to be sustained. Instead the household sector is likely to either use savings to pay for higher fuel prices or reduce other types of purchases in order to meet their budget. The magnitude of the increase over the last winter is large enough to be noticed by consumers but is unlikely to cause radical changes in consumer behavior. Of course if energy is a larger portion of any individual household’s budget, the effect will be more pronounced.
8 The prices used for coal are based on the delivered utility price per ton and not the price for central Appalachian coal used in the introduction. Given that coal is used almost exclusively for electricity generation, this represents a fair estimate of the cost paid by utilities.
9 North America has only 4.2 percent of the proved natural gas reserves in the world, it produces 21 percent of the world’s supply and accounts for 30 percent of the world’s demand. In contrast the Middle East has nearly 41 percent of the world’s proved supply and Europe and Eurasia has 35 percent.
10 A Btu is defined as the amount of heat required to raise the temperature of one pound avoirdupois of water by one degree Fahrenheit. Normalizing the price by Btu allows a comparison of the resources needed to create this amount of heat and can be used as a rough proxy for the heating price of a particular fuel stock.
11 Briefing by Finley (2005).
12 A report by the consulting firm Accenture and Cambridge Energy Research Associates issued in 2003 analyzed the impact of market volatility on 16 energy companies and found that volatility was preventing increased investment in energy assets. (See Accenture and Cambridge Energy Research Associates, 2003.) However economists have been more ambivalent about the impact of volatility on investment. The key determining factor influencing the relationship between volatility and investment include the life of the investment, whether the investment is reversible, the nature of competition that firms in the industry are facing, and the relative risk aversion of firms in the industry. Given these factors, both empirical work and theoretical work come to widely differing conclusions about whether volatility helps or hinders investment. However, it would appear that most energy companies do have the profile (long-lived assets that tend to be irreversible once started in an industry known for risk aversion) that would suggest that volatility would impede investment. For a more complete discussion see, Pindyck (1988), pp. 969–985. For an interesting empirical study, see Bell and Campa (1997), pp. 79–88.
13 A measure of the coldness of the weather experienced, based on the extent to which the daily mean temperature falls below a reference temperature, usually 65 degrees Fahrenheit. For example, on a day when the mean outdoor dry-bulb temperature is 35 degrees Fahrenheit, there would be 30 degree days experienced. A daily mean temperature usually represents the sum of the high and low readings divided by two. A form of degree day used to estimate energy requirements for air conditioning or refrigeration. Typically, cooling degree days are calculated as how much warmer the mean temperature at a location is than 65 degrees Fahrenheit on a given day. For example, if a location experiences a mean temperature of 75 degrees Fahrenheit on a certain day, there were 10 CDD (cooling degree days) that day because 75 – 65 = 10.
14 Diaz and Quayle (1980) found that the correlation between energy use and heating degree days was as high as .97 at the household level. Energy consumption increases as the number of heating and cooling days increase in a highly related relationship. See Diaz and Quayle (1980), pp. 241–246.
REFERENCES


CALL FOR PAPERS AND SAVE THE DATE

May 17–19, 2006

42nd ANNUAL CONFERENCE ON BANK STRUCTURE AND COMPETITION
FEDERAL RESERVE BANK OF CHICAGO

Innovations in Real Estate Markets: Risks, Rewards, and the Role of Regulation

The Federal Reserve Bank of Chicago invites the submission of research- and policy-oriented papers for the 42nd annual Conference on Bank Structure and Competition to be held May 17–19, 2006, at the Westin Hotel in Chicago. Since its inception, the conference has fostered an ongoing dialogue on current public policy issues affecting the financial services industry. We welcome submissions of high-quality research papers on all topics related to financial services and their regulation. As in past years, the program will highlight the conference theme, but will also feature numerous sessions on topics unrelated to the conference theme.

This year the conference theme will explore the changing economics of real estate markets. Over the past decade, the U.S. real estate sector has been an important engine for macroeconomic growth with innovations in home mortgage finance playing a central role. New mortgage production methods based on credit scoring, asset securitization, and mortgage brokerage have greatly enhanced household access to home mortgage credit. Coupled with historically low interest rates, these innovations have driven home ownership rates to all-time highs. Adjustable rate mortgage products have supplanted traditional fixed rate mortgages in many local markets, allowing households to leverage the long decline in mortgage rates into larger homes and/or greater disposable income. Similarly, more and cheaper refinancing options—especially cash-out options—allow consumers to unlock the equity value stored in their homes. These developments supported consumer spending growth, which in turn dampened the most recent recession and fueled the recovery that followed.

However, some observers see dark economic clouds on the housing horizon. On the one hand, rising home prices have been a boon to personal wealth, offsetting historically low household savings rates and forming a base for consumer spending. But on the other hand, the rapid increases in home prices in coastal markets look like asset price bubbles to some; if and when these bubbles burst, the negative effects from falling home prices on wealth could slow economic growth. Speculative buying in especially hot housing markets could exacerbate these outcomes; if and when economic growth slows, pressure to sell these properties would further depress prices. Other observers worry about households’ increased exposures to interest rate risk; if and when interest rates move...
up, delinquent payments and loan defaults could increase among homeowners, especially those that relied on adjustable rate mortgages to afford their homes. If these potential events were to unfold together, a worst-case scenario could put stress on the huge mortgage portfolios held by the housing government-sponsored enterprises, with potential systemic effects in financial markets.

Innovations in real estate markets also create the potential for social change. The evolution in home mortgage finance has made home ownership a reality for more than two-thirds of U.S. households—and many social scientists believe that home ownership is a stabilizing social and economic force. But along the way, home mortgage finance has become more complicated. Today, the majority of newly originated mortgages contain some “exotic” feature such as adjustable rates, interest only payments, or optional negative amortization. When used appropriately, these innovations work to the advantage of the consumer by making home ownership more affordable. But the complexity of these arrangements—especially when they are coupled with low down payments—can tempt naïve or overly optimistic home buyers into assuming unaffordably high debt loads. A debate is emerging about the role of public policy in these cases, ranging from mandatory measures to inform home buyers to better definition and policing of predatory lending practices.

In addition to the above issues directly related to the 2006 conference theme, the program committee is interested in evaluating submissions on the following topics:

- Competitive strategies of financial institutions;
- The Basel II Capital Accord;
- Financial industry consolidation;
- Payments innovations;
- Fair lending and the Community Reinvestment Act;
- Small business finance;
- Deposit insurance and safety net reform;
- Measuring, monitoring, and managing bank risk;
- The viability and future role of community banks; and
- Restructuring of financial regulatory agencies.

If you would like to present a paper at the conference, please submit a completed paper or a detailed abstract (the more complete, the better) via email to Regina Langston at rlangston@frbchi.org by December 10, 2005. Please include your name, address, affiliation, telephone number, and email address, and those of any co-authors.

For additional information contact:
Douglas Evanoff at 312-322-5814,
Robert DeYoung at 312-322-5396, or
Regina Langston at 312-322-5641.
Shopping without cash: The emergence of the e-purse

Carol L. Clark

Introduction and summary

During the 1990s, some payment analysts suggested that smart cards with e-purse applications could be a promising new payment option for certain types of transactions. An e-purse is a stored-value payment device that offers the following features to the consumer: It holds electronic monetary value that substitutes for cash; it does not require online authorization; it records the value of each purchase on the card rather than a central computer server; and it can be exchanged for goods and services from various merchants. The device is generally stored on a computer chip, which can reside on any one of a number of items most consumers already carry, such as a payment card, mobile phone, key chain, or even a watch. When the consumer makes a purchase, monetary value is deducted from the microchip on the card.

The key difference between a stored-value smart card and debit, credit, payroll, and gift cards is that value is stored directly on the smart card rather than stored in an account on a central computer server, and therefore, transactions are processed offline between the smart card and the card reader at the point of sale (POS). In contrast, debit, credit, payroll, and gift cards in the United States are offered on magnetic stripe cards, and payment involves an online authorization that requires a real-time connection with a central computer. The purchase is approved or declined through the authorization process, which checks whether there is sufficient value in the account for debit, payroll, and gift card transactions and whether the credit limit has not been exceeded for credit card transactions. The authorization process may also check whether the card is fraudulent or stolen.

Some payment analysts predicted that smart cards could lead to a cashless society, one in which e-purses would replace cash and coins for low-value payments. As we know, this hasn’t happened. Although a number of e-purse programs have been implemented around the world, these programs have experienced varying degrees of success, and many have failed outright. Smart card adoption in the United States has been slower than in the rest of the world. Many analysts argue that this is partly because the U.S. already has an advanced telecommunications infrastructure that can verify magnetic stripe card transactions quickly and cheaply online. This results in relatively low fraud levels and relatively high levels of satisfaction among businesses and consumers with the current systems. If this is true, then smart card applications may offer more value in other parts of the world with less highly developed telecommunications infrastructures and higher incidences of fraud in existing payments networks.

In this article, I review six e-purse smart card programs in Hong Kong (one) and the United States (five). I chose these two regions because Hong Kong has one of the most highly successful e-purse programs, the Octopus card, and the United States has implemented a number of e-purse programs, some of which have been more widely adopted than others. I find that the most successful among these programs tend to have the following characteristics: a captive audience that drives critical mass, such as those found in the transportation industry or government sector; an affordable cost structure relative to other payment instruments; compelling incentives to consumers and merchants; and a technology that is well tested and addresses standards issues before the rollout.

Carol L. Clark is a payments research manager at the Federal Reserve Bank of Chicago. The author gratefully acknowledges the assistance of Erin Davis, Juan A. De Jesus, David Doyle, Tamara Kidder, Graham Mackenzie, John Scaggs, Barbara Straw, Eric Tai, and Joey Wong in the completion of this study and the helpful comments on previous drafts by Sujit Chakravorti, Geoffrey Gerdes, Richard Porter, Tara Rice, and Leo Van Hove.
Below, I survey the theoretical framework of previous smart card studies, provide an overview of the payments environment in Hong Kong and the United States, and analyze six e-purse programs in these two regions and the factors that contributed to their success or failure. Then, I discuss the implications of my findings for future e-purse programs.

**Literature review**

One of the greatest challenges in the adoption of a new payment device is establishing a critical mass of users. Regardless of the type of technology used, consumers are reluctant to use a new payment instrument if few merchants accept it, and merchants will refuse to accept the device because the cost of installing and maintaining the supporting technology infrastructure, like card readers, may be prohibitive, unless enough consumers want to use it. New payment mechanisms gain momentum when enough people use them, which leads to widespread acceptance by the merchant community. Critical mass, however, is not only related to the number of users but also to the actual levels of usage because the program’s profitability is generally dependent on high transaction volumes (Goldfinger, 1998). As Rochet and Tirole (2003) observe, merchants cannot benefit much from consumers that hold a payment card but use it only sporadically. The more frequently the card is used, the more valuable it becomes to consumers and merchants. Therefore, frequent use is one of the keys to a successful e-purse program.

Goldfinger (1998) estimates that a critical mass of one million users was needed for a smart card program to attain profitability due to the large fixed costs of the infrastructure, although these costs have likely fallen in recent years. To achieve this, Goldfinger argues that program promoters have to be able to orchestrate a large-scale deployment and initiate a migration/switching process from the existing payment system to the smart card system. He takes the view that the benefits that smart cards provide cannot be fully realized if there is an alternative payment infrastructure present. While this is certainly not the case for mature payment infrastructures—cash, checks, debit cards, and credit cards coexist at most retailers—there may be some validity to this argument in the case of an emerging payment instrument like an e-purse.

In another study, Van Hove (2004) examines data on 16 e-purse systems in Europe. Van Hove finds that successful programs are in countries that are relatively small geographically or have phased introductions; that have online debit card systems that are fairly popular or cannot be used for low-value payments; that have stakeholders that quickly agree on a common solution so there are no incompatibility problems; that have major banks committed to and participating in the program; and that have support from key players that operate and support one or more of the following: public telephones, parking meters, vending machines, or public transportation.

Chakravorti (2004) finds three other necessary conditions for a viable new payment instrument: There must be benefits that are not provided in existing payment instruments for at least certain transactions; consumers and merchants must be convinced of these benefits and, possibly, provided with incentives to change their behavior; and the new system must be perceived as secure, with adequate measures against credit risk and fraud.

**Payments environment in Hong Kong and the United States**

As I explained in the introduction, I am interested in comparing programs in Hong Kong and the United States because Hong Kong has one of the most highly successful e-purse programs, the Octopus card, and the United States has implemented a number of e-purse programs with varying degrees of success. As figure 1 shows, Dove Consulting (2003) reported that in 2003 electronic payments surpassed other types of payments for in-store purchases for the first time in the United States. However, cash was still the most popular payment vehicle.

Cash is used even more widely in Hong Kong. Eric Tai, chief executive officer of Octopus Cards Ltd., indicates that Hong Kong residents use coins

![Figure 1: How U.S. consumers made in-store payments, 2003](source: dove consulting (2003)).
and currency 50 percent of the time. Checks are used for retail transactions, where credit and debit cards are not accepted, and credit cards have become increasingly popular, with over nine million in circulation in 2001 (Bank for International Settlements, Committee on Payment and Settlement Systems, 2003). Interestingly, however, in Hong Kong, e-purse transactions are now growing faster than either debit or credit card transactions—Euromonitor International (2004) reports that they increased by 8 percent in 2003, compared with 7 percent growth in debit card transactions and 2 percent growth in credit card transactions.

While the United States and Hong Kong have each implemented a number of e-purse programs, only Hong Kong’s Octopus card, which began in the niche transportation industry and extended outward to retailers, has been widely adopted by consumers and a diverse number of merchants. More than 95 percent of Hong Kong’s residents aged 15–65 carry the card. Over 50,000 smart card readers accept Octopus at public transportation terminals, convenience stores, fast food chains, leisure facilities, parking meters and garages, pay phones, personal care stores, photo booths, photocopiers, school snack shops, supermarkets, taxis, and vending machines (Tai, 2005). In August 2005, Octopus announced an apparel retailer will accept the card at its Hong Kong locations. Some e-purse programs in the United States that began in niche markets are currently successful, but on a much smaller scale.

Octopus processes over nine million transactions each day with an average daily transaction value of about HK$65 million (US$8.3 million) amounting to about 2 percent of Hong Kong’s gross domestic product (GDP) in 2003 (U.S. Department of State, Bureau of East Asian and Pacific Affairs, 2004). Retail purchases in Hong Kong using the Octopus card grew from 5 percent in January 2002 (Trintech Group Plc, 2003) to 17 percent of total transactions in August 2005 (Wong, 2005). With about US$1.4 million in average daily retail transactions, Octopus takes in more in a single day than the widely reported Mondex and Visa Cash trial in New York City did during the entire 15-month program.

As I mentioned earlier, most payment analysts agree that smart card adoption in the United States has been slower than in the rest of the world because the United States has an advanced telecommunications infrastructure that can verify magnetic stripe credit and debit card transactions quickly and cheaply online. This results in relatively low fraud levels and relatively high levels of satisfaction among businesses and consumers with the current systems. Smart card applications may offer more value in other parts of the world with less highly developed telecommunications infrastructures and higher fraud incidences.

The business case for smart cards in the United States also depends on a number of other factors. There are issues related to who would pay for the extra chip on the card and to what fees merchants would pay on a per transaction basis. In Hong Kong, merchants appear to be paying lower rates on Octopus transactions than on credit card transactions.

There are also differences in the technology used for stored-value cards in the two regions. Octopus provides e-purse capabilities on a contactless smart card, which means the card does not have to be inserted into a card reader like credit or debit cards. Instead, it is held close to the reader and payment is registered in 0.3 seconds. Meanwhile, Duetto cards offered by the coffee chain company Starbucks, payroll cards that are used instead of direct deposit or paychecks by some firms to deliver an employee’s pay, and gift cards offered by various retailers in the United States provide stored-value capabilities on magnetic stripe cards. There are two ostensible reasons for using magnetic stripe cards rather than contactless smart cards in the United States: the cost of equipping stores with chip reading terminals and the desire to include Visa, MasterCard, or private label branding since these cards are processed by online readers. In addition, some payment providers in the United States offer contactless smart cards but link purchases to credit card or debit card accounts rather than to an e-purse—examples include Exxon-Mobil’s SpeedPass, Bank of America’s QuickWave, and MasterCard’s PayPass. In an interesting development, in December 2004, the Washington Metropolitan Area Transit Authority began piloting 20,000 MasterCard branded magnetic stripe cards that also contain a stored-value chip for transportation (Garback, 2005).

Case studies

I examine six e-purse case studies that began in “closed-loop” environments in Hong Kong and the United States, meaning they were offered to what one might call a captive audience, such as one found in a military facility or university campus. The e-purse programs that were tested in open-loop environments in these two regions have failed outright, such as the Mondex and Visa Cash trial in New York City cited previously. I chose the case studies to represent a cross section of industries that have implemented e-purse programs in recent years: transportation, government, and higher education. The Octopus card’s e-purse transaction volumes and values are among the highest in the world. The Ohio Electronic Benefit Transfer program, which has higher transaction values and volumes than
Octopus, is the largest smart card program for administering food stamps in the United States. The University of Michigan Mcard represents one of the largest university deployments of an e-purse in the United States. The University of Central Florida UCF Card is one of the few campus e-purse programs still in operation. The Navy Cash card and the EagleCash card programs are two of three smart card programs administered by the U.S. Department of the Treasury for the U.S. Armed Forces. A synopsis of the six programs, as well as a detailed discussion on how each card works, is included in the appendix.

**Octopus card**

The Octopus card began in the niche transportation industry when Hong Kong’s five leading companies for trains, buses, ferries, and subways formed a joint venture in 1994 to oversee the implementation of a smart card system. After three years of development and trials, they launched Octopus in 1997. In 1999, 7-Eleven stores in Hong Kong became the first locations outside the mass transit system where riders could add value to cards. The convenience store chain liked the speed and ease of the contactless technology so much that it installed readers in its stores in the following year so that consumers could pay for goods using Octopus. In time, consumers began to press other retailers to accept the card as well (Ramstad, 2004).

A number of factors were crucial to the success of the Octopus card: the support of five transportation companies; the interoperability of the system; the manner in which critical mass was established by leveraging the captive and niche transportation industry; the reliable technology; and the compelling incentives offered to consumers and merchants.

**Factors influencing success**

Octopus has the support of Hong Kong’s five major transportation companies. Although some of these companies compete directly for riders, the savings they achieved by implementing a shared smart card system appear to have outweighed any competition concerns (Poon and Chau, 2001). This also implies that the profit-sharing scheme the transportation companies worked out is equitable enough to induce cooperation. For consumers, the development of a single interoperable system means they can access any public transportation in Hong Kong with the same card. In contrast, 40 miles from Hong Kong in Macau, two bus companies launched separate incompatible e-ticket systems that failed to reach critical mass because traveling in the area typically requires a combination of buses and most people were not willing to carry two different cards (Uzureau, 2003).

Octopus has also been free of technology-related problems, unlike several smart card programs that have had trials in the United States. Very few failures of the Octopus card were reported during the first month of operation. On average, station personnel needed to resolve problems in only one out of every 11,000 journeys (Wynne, 1998).7

Octopus also uses radio frequency identification (RFID) technology, which allows commuters to wave their card (or a purse or wallet containing the card) within 4 inches of the reader at the ticket barrier to register payment within 0.3 seconds (BusinessWorld Publishing Corporation, 2002). Thus, an Octopus card transaction takes less time than a cash transaction in which one may have to wait for change, and takes significantly less time than the typical credit or debit card transaction in which magnetic stripe technology is used. Moreover, the durable smart cards have a potential life span of about 100,000 transactions (Tai, 2005). And Octopus’s functionality has been embodied in a variety of forms, including key chains, mobile phones, and watches.

What about incentives? Initially, Octopus offered consumers a 10 percent savings and a 100 percent satisfaction guarantee to increase adoption in the transportation sector and to remove uncertainty about the new technology (Tai, 2005). These incentives, along with the simplicity, speed, and convenience of the system’s technology, resulted in over three million cards being issued during the first three months and established a critical mass of smart card users who were familiar with RFID technology.

Metro and rail transportation operators offer multiple ride tickets on the Octopus card and single ride tickets on magnetic stripe cards (Wong, 2005). This is significant because over 70 percent of Hong Kong residents use some form of public transportation each day (Poon and Chau, 2001) and are more likely to use the multiple ride tickets offered by Octopus. Tai (2005) reports that constraining multiple ride tickets to Octopus cards elicited little consumer dissatisfaction. Metro and rail transportation operators provide discounts to Octopus cards over single ticket cards; the discounts vary according to the distance traveled. Smart card adoption for metro riders is 90 percent and for rail commuters over 80 percent (Wong, 2005).

Transportation operators for buses, minibuses, and ferries accept coins or Octopus cards, and fares are the same for each payment method. Octopus card adoption on these transportation lines is somewhat lower compared with the metro and rail lines—70 percent for ferry lines, about 80 percent for minibuses, and over 80 percent for buses. Although buses, minibuses,
and ferries do not consistently offer a discount to Octopus cardholders as do the metro and rail lines, they do sometimes launch promotional campaigns that offer discounts to Octopus cardholders only.

Once a critical mass of smart card users was established in the transportation industry, the proven technology was used to branch out into the retail market, where consumers were offered a number of benefits that helped foster adoption. Octopus is a single convenient, multipurpose card that speeds retail transactions and replaces cash for small purchases. In contrast to other e-purse programs, Octopus actually allows cardholders to make purchases up to a negative value of HK$35 (US$4), so long as the card contains a positive value of HK$0.01 before the purchase. Once the card has a negative value, it must be reloaded before it is used again. Octopus recovers the negative balance through the deposit and purchase price of the cards. For a detailed discussion of the types of Octopus cards, deposit amounts, and card costs, see the appendix.

Merchants also enjoy a number of benefits. Octopus reduces cash handling and in-store queues, and increases customer loyalty by allowing merchants to offer ad hoc discounts to customers using the card. It is difficult to determine the cost to retailers of accepting the card, since data on hardware costs and merchant fees are confidential. The World Bank’s website indicates that Octopus has a two-part transaction fee. There is a HK$0.02 charge for every transaction to cover the costs of technical support, computer operations, and replacement cards and a 0.75 percent charge on the transaction value to cover card-control operations, legal, marketing, and depreciation costs. Therefore, a HK$10 transaction would include a fee of HK$0.02 plus HK$0.075, or HK$0.095 (Rebelo, 1999). However, Octopus Cards Ltd. has indicated that these transaction charges vary depending on merchant volume (Cheng, 2004).

Despite the uncertainty about exact costs, it appears likely that retailers in Hong Kong benefit from lower transaction fees for the Octopus card relative to transaction fees for credit cards, which vary from 2 percent to 4.5 percent (Morgan and Snee, 1997). Although new locations like McDonald’s are accepting Octopus (Tai, 2005), some merchants still find Octopus fees to be too expensive. In CardTechnology, Balaban (2005) reports that a few retailers like Starbucks have reduced the number of outlets that accept Octopus.

Ohio EBT program

In the United States, the U.S. Department of Agriculture, Food and Nutrition Service (FNS) has shifted qualified low-income families from paper food stamp coupons to electronic benefit transfer (EBT) cards. The EBT program was designed to reduce fraud, to eliminate the cumbersome manual processes associated with issuing and redeeming paper food stamps, and to lessen the stigma associated with being a traditional food stamp recipient. In 2003, 9.1 million U.S. households redeemed an average of $1.7 billion in food stamps every month using EBT cards. To reduce fraud, the system creates an electronic record of each transaction that can help identify where food stamps are trafficked or exchanged illegally (U.S. Department of Agriculture, Food and Nutrition Service, 2004).

States have taken different approaches to administering the EBT program. Forty-eight states have implemented magnetic stripe systems that require online authorization from a host computer that keeps track of value. Two states, Ohio and Wyoming, use offline smart card systems that store value on a computer chip resident on the card. The state of Ohio has announced, however, that it is discontinuing its smart card program, Direction Card, which has been in place since 1996, and is seeking bids for an online system (Welsh-Huggins, 2003).

Factors influencing failure

John Scaggs (2005), Ohio’s EBT project director, indicated that the decision to discontinue Ohio’s offline system was based on cost, as well as on the failure of credit card companies to build a smart card infrastructure, which had been anticipated when the program was implemented in the mid-1990s. The online system will be installed no later than June 2006.

The decision to discontinue the program followed a 2002 study by Abt Associates, Inc. (2002) that compared Ohio’s program with the findings of the three most recent EBT system evaluations. These included the online system in the state of Maryland; the offline pilot in Dayton, Ohio, on which the Direction Card system was later built; and the offline system in the state of Wyoming. The study found Direction Card was more expensive than Maryland’s online system, but less expensive than the offline systems in Dayton, Ohio, and Wyoming. Abt Associates estimated that the total operational costs of the Direction Card system were 56 percent higher than Maryland’s system due to more expensive hardware, software, and local agency costs. The Direction Card was 29 percent less expensive than the Dayton pilot because of the larger scale of the Direction Card program, the lower costs of building the Direction Card system upon the Dayton pilot, and the decreased technology costs resulting from technological developments that emerged after the Dayton pilot was deployed. The Direction Card
was 43 percent less expensive than Wyoming's program due to lower local, state, customer service, data center hardware, software, POS, and card costs (Abt Associates, Inc., 2002). The Abt Associates study also reviewed advantages and disadvantages of the Direction Card from the perspective of retailers. Ohio's merchants received free terminal installation, initial user training, and maintenance; however, they never found a cost-effective way to integrate the offline system into their existing online POS devices. Therefore, Ohio retailers bore the ongoing costs of training staff to use the separate terminal, not to mention lost counter space. There were also differences in costs related to equipping store lanes with EBT POS devices. Even though the number of POS devices given to large retailers by Ohio was more generous than the FNS mandated, large retailers did not have enough terminals to equip every checkout lane in the store. To do so, they would have to pay for extra POS devices. In contrast, online merchants that integrated EBT transactions into existing POS devices could service EBT customers in any lane. However, most online systems did not pay the costs of integrating cards into existing POS devices, which then shifted the costs of doing so to the merchant.

Those online merchants that did not integrate EBT transactions into existing POS devices either used the state's allotment of EBT terminals or paid for extra equipment, but online terminals were less expensive than offline terminals (Abt Associates, Inc., 2002). The Abt Associates study also compared the experiences of Ohio's EBT cardholders with those of EBT recipients accessing online systems. While Ohio recipients had higher levels of service due to hands-on training at Direction Card system offices, this specialized training also required extra time from the cardholder and sometimes necessitated an added trip to the local office. In contrast, states with online systems piggybacked on the widespread use of magnetic stripe cards for other applications, as well as typically providing cards, training materials, and personal identification numbers (PINs) by mail. There were also differences in loading value onto the card. Ohio EBT recipients are required to load their benefits at any one of three stores of their choosing or at their local food stamp office. In contrast, EBT recipients with magnetic stripe cards do not have to load value on the card at any specific location because value is stored on the central computer server (Abt Associates, Inc., 2002).

In addition, EBT smart card recipients may have experienced more confusion about the current value on their cards. Almost 90 percent of the Direction Card calls to customer service centers were to check the card balance. Similarly, the majority of customer service calls for online systems were to check account balances. The report indicates, however, that there may be an additional reason for balance inquiry calls for offline cards. The Direction Card system deducts the purchase amount from the chip on the card at the time of purchase. Information about the transaction is sent to the central computer server via batch processing at the end of the day when the beneficiary's account information is updated. In contrast, online systems verify transactions real-time against the central computer server, and balances are updated immediately. Thus, offline card users may have been confused by the balance information on the audio response unit, which obtains information from the central computer server on a lag basis, compared with their knowledge of the available balance based on known card transactions and expected benefits (Abt Associates, Inc., 2002).

The Abt Associates report also considered the transferability and adaptability of Ohio's Direction Card to food stamp programs in other states. The main obstacle cited was the need to build a system from scratch, since few retailers have POS devices capable of reading smart cards because consumer demand for these cards has not reached a critical mass. Moreover, EBT recipients outside Wyoming and Ohio benefit from the interoperability of online systems, allowing them to access benefits in 48 states. Ohio and Wyoming recipients can only use smart cards in their own respective states, unless an out-of-state store is specially equipped to accept them.

Campus smart cards

Numerous e-purse programs have been implemented in closed-loop college and university environments for a variety of reasons: Students are open to new technologies; universities are able to implement more secure IDs that are not as easily duplicated as magnetic stripe cards; and schools are able to reduce administrative costs and to generate transaction fee income. Despite these benefits, most campus smart card trials have failed. In University Business magazine, Villano (2004) reports that of the approximately 50 schools in the United States that implemented smart card programs from 1997 to 2002, only a handful are still using them and relatively few take full advantage of the capabilities the technology provides. Failures are attributed to the high costs of offline systems compared with those of online systems, lack of interoperability, and delays in batch processing that mean card balances may not be updated for up to 24 hours.
University of Michigan

One of the largest e-purse deployments at a campus in the United States began at the University of Michigan in 1995 in response to students' and merchants' requests that the school's Entrée Plus system be extended off campus (Mayer, 1996). Entrée Plus was a funds pool into which parents put money at the beginning of the year for meals at residence halls and snack bars, as well as for on-campus purchases at vending machines, bookstores, and laundry facilities (Mitchell, 1998). As it turned out, it was not feasible to expand the Entrée Plus system off campus, so the university developed a proprietary offline POS network called the Mcard (Mayer, 1996).

In June 2001, the university announced that smart cards would be gradually discontinued and replaced with magnetic stripe cards (Avisian, Inc., 2002). A number of factors influenced the failure of the offline system, including technology problems; the overall cost relative to online systems; the lack of a critical mass of users and merchants; confusion related to concurrent programs on the card; and the apparent lack of a business case at the outset.

Factors influencing failure

In terms of technical problems, outdoor card readers did not function well in cold weather, and transaction times took longer than cash (Mitchell and O'Brien, 1999). In addition, the chip on the Mcard malfunctioned nearly one-quarter of the time when it was first implemented (Doyle, 2005). Some cashiers had not been properly trained, which resulted in delays, and students became frustrated and eventually mistrustful of the card (Michigan Daily, 1996).

Over time, the school found that the Mcard was more useful as a coin substitute than as a paper currency replacement. As such, the school required technology that would support pay phone, parking meter, and central parking facility transactions, but these were beyond the capability of the system (Doyle, 2005). Faced with an obsolete technology, the university sought bids from Visa, Mondex, and Proton to update the program. However, costs were considered too high, and the school announced it would revert to an online system (Kuykendall, 2001).

Moreover, usage of the Mcard was lower than expected. Over a year after its implementation, transaction volumes were 20 percent less than anticipated and dollar values 30 percent to 35 percent below target (Chakravorty, 1996). Although usage rates would generally be lower at the beginning of any new payment technology, the Mcard illustrates a classic dilemma: Cardholders wanted more merchants to accept the card, while merchants wanted more cardholders to use it.

It was also hoped that the Mcard program would provide students and staff with a convenient payment tool that would generate revenues for the school. The Mcard included a microchip for payments, a bar code for checking out library books, and a magnetic stripe that functioned as a debit card as well as provided building access. Phone card functionality was also included on the card (Mitchell, 1998). The University of Michigan received part of the ATM transaction fees and calling card revenues in addition to 50 percent of the merchant transaction fees paid to the school's bank, First of America (Doyle, 2005). However, students and parents were confused by the many concurrent programs that operated separately on the card. In addition, students preferred the Entrée Plus program to the e-purse program. With Entrée Plus, parents could deposit funds to a university account that students could draw from to make on-campus purchases, whereas with the e-purse, students typically used their own money to load value onto the chip (Mitchell and O'Brien, 1999).

Moreover, the on-campus rollout of the Mcard was not well coordinated, and merchant fees were viewed as high. Only one of the campus's 22 libraries accepted the card for photocopies because most had already implemented their own copy systems prior to Mcard's launch (Doyle, 2005). Some other campus locations did not accept the Mcard, and the number of purchases with the Mcard was lower than with credit cards. On-campus merchants initially paid fees of 4 percent for transactions on the chip and BankStripe (a closed-loop debit card linked to a First of America checking account), which were later lowered to 2.8 percent, while off-campus merchants consistently paid 4 percent (Doyle, 2005). Merchant interest in the program was also low due to the high cost of the card reader, $900, resulting from the lack of competition from other vendors (Gale Group, Inc., 1998).

Finally, it is unclear whether a business case for a smart card over magnetic stripe technology was fully developed. The Mcard program was originally piloted as a magnetic-stripe ID card. When the supplier that provided the cards and readers was purchased by a leading smart card vendor a few months after the pilot began, the university's bank suggested that a microchip be added to the magnetic stripe card to facilitate faster on- and off-campus payments and to provide students with a cash substitute for pizza deliveries to dorms and for meals in restaurants (Mitchell, 1998).
Joining wide program

United Federal operations

cards of that card or card of reports in-house student and student interest in expanding the program off campus.

Factors influencing success

Tamara Kidder (2005), the UCF Card manager, reports that the university now performs its own transaction processing, which is quicker than the processing offered by their initial vendor. Moreover, the number of students and staff using the UCF Card increased from 20 percent in 2003 to 50 percent in the summer of 2004 as a result of the computer lab requiring the card for print copies. Students also enjoy the security of two PIN-protected e-purses; if the UCF Card is lost or stolen, these funds can be transferred to a new card.

Similar to the evolution of the Octopus card, the expansion of the UCF Card off campus was driven by both student and merchant demand. Currently, the card is used to purchase on-campus pizza deliveries and off-campus meals at six fast food and other restaurants. Twenty additional merchants are interested in joining the program (Kidder, 2005), which suggests that the 3 percent merchant discount fee is not viewed as a barrier.

Military smart cards

One other sector where e-purse applications have been developed and implemented successfully in the United States, albeit in a fairly narrow range of cash management applications, is the U.S. government. Since the inception of its stored-value card (SVC) program in 1997, the U.S. Department of the Treasury has issued more than 1.4 million smart cards that support specific business processes within each branch of the U.S. Armed Forces (Mackenzie, 2004). The program aims to end the float loss associated with the more than $2 billion in coin and currency in circulation on military bases, ships, and other locations worldwide—and the associated cost of securing, transporting, and accounting for cash held outside the Treasury. The cards also eliminate the manually intensive back-end operations necessary to support scrip, vouchers, meal tickets, money orders, traveler’s checks, and other paper payment mechanisms used in closed government environments. Two types of SVC cards currently in use are Navy Cash™ and EagleCash.

Navy Cash™ card

Navy Cash™ reduces the cash handling and fiduciary costs associated with safeguarding and storing large amounts of paper currency and coins on board ships by combining two technologies on the same card: smart card chip and magnetic stripe. The e-purse on the chip is used for onboard purchases, and the magnetic stripe is used for purchases during shore leave. On a recent voyage of the USS Harry S. Truman, which has a crew of 5,000, there were over US$1 million in sales using the Navy Cash™ card (Gosnell, 2004). As of September 30, 2005, 66 ships have been deployed with Navy Cash™, and a total of 160 are scheduled to convert to the system by 2008 (Straw, 2005a).

Factors influencing success

A number of factors are influencing the current success of the Navy Cash™ program, including cost savings for the Navy, high customer satisfaction among crew members, and powerful incentives that drive critical mass.

Navy Cash™ reduces labor-intensive cash handling as well as lowers fiduciary and reporting costs by decreasing the need for paper currency and coins on board ships by 50 percent to 60 percent. In addition, a “cashless” shipboard environment allows more time for crew members to focus on core mission functions. Navy Cash™ also provides crew members with greater services and security, and supports access to home bank and credit union accounts. The Navy Cash™ system has been taken successfully around the world with a high rate of customer satisfaction (Straw, 2005b).

The Navy is very motivated to cut costs, and provides cards free of charge to crew members. While adoption of the program by servicemen and service-women is voluntary, those crew members who choose to forgo the card cannot make onboard purchases at retail POS terminals or at vending machines. Instead, they only receive standard food and supplies issued by the Navy, as no cash is accepted on the ship (Rivers, 2004).

EagleCash

A second military stored-value smart card called EagleCash began in 1999. It is used by U.S. Army and civilian workers in parts of the world with weak telecommunications and banking infrastructures. Personnel use the reloadable card on bases to purchase goods and services as well as foreign currency. First
deployed in Bosnia, the EagleCash smart card is also used in Afghanistan, Honduras, Kosovo, Macedonia, Qatar, and Uzbekistan. Future deployments may include military bases in Iraq, Kuwait, Saudi Arabia, and the Sinai (De Jesus, 2005).

Factors influencing success

The current success of the EagleCash program can be attributed to a number of factors. Staff is able to pay for goods in countries where banking and telecommunications infrastructures are weak and benefit from faster checkout lines at on-base POS terminals, where checkout times decreased by an average of 45 seconds (Federal Document Clearing House, Inc., 2000). The military benefits from a reduction in U.S. currency in hostile areas, which could potentially be used to fund terrorism due to the stronger market value of the U.S. dollar in these regions (Thompson, 2004). The risk of counterfeiting and black marketing is also diminished. In the Balkans, EagleCash lowered U.S. currency in circulation by $160 million (De Jesus, 2005).

Implications for future e-purse programs

A review of six e-purse programs in Hong Kong and the United States reveals that the successful programs operate in captive markets and sometimes require use of the card to establish a critical mass of users as seen in the Direction Card, Navy CashTM, UCF Card, and Octopus programs. For example, the Direction Card program required EBT recipients to migrate from paper food stamps to smart cards. Navy CashTM is almost mandatory, since service personnel are unable to make onboard purchases without the card. University of Central Florida students are required to use the UCF Card in the school’s computer lab, and this requirement resulted in an increase in the number of students and staff using the card from 20 percent to 50 percent. Octopus leveraged the transportation industry in Hong Kong, where over 70 percent of its commuters take public transportation every day. Commuters who travel via rail and metro can only purchase multiple ride tickets using the Octopus card. Of course, required use of smart cards is possible in closed-loop environments; however, mandatory usage of smart cards has little chance of successful implementation in open-loop environments where consumers are accustomed to payment choices.

Octopus’s initial focus on the transportation sector also supports Van Hove’s (2004) observation that successful smart card programs typically receive support from key players in at least one of the following: public transportation, public telephones, parking meters, or vending machines. Public transportation proved to be the conduit for establishing critical mass in Hong Kong. Is such a system feasible in the United States? Greg Garback (2005), executive officer of the Finance Department of the Washington Metropolitan Area Transit Authority, indicates that U.S. transportation companies are investing $1 billion to improve their payment infrastructures through the use of contactless smart card technology. Cities with multimodal public transportation (trains, buses, subways, or ferries) like Chicago, Seattle, San Diego, Los Angeles, Boston, Houston, Atlanta, and San Francisco have installed contactless smart card systems. These regional transportation systems may one day lead to a single transit card that can be used in multiple cities.

The Netherlands is currently installing such a cross-country transportation system, which should be completed by 2006 (ASK, 2004). Octopus is providing back office support and operational expertise on the project (Tai, 2005). However, compared with the United States, both the Netherlands and Hong Kong are extremely small geographic areas, which certainly facilitate the implementation of a systemwide or countrywide e-purse strategy, as Van Hove (2004) observes. Furthermore, Hong Kong residents use public transportation daily. By comparison, the U.S. Census Bureau (2004) reports that there is only one U.S. city where the majority of the population (55 percent) takes public transportation—New York City. Other cities with a population of 250,000 or more have a smaller percentage of the population that takes public transportation each day.

Because the vast majority of commuters in the United States as a whole (77 percent) drive to work alone, some payment analysts argue that the RFID technology currently used in tollway systems in some metropolitan areas may be expanded to the retail sector in the future.15 In fact, the E-ZPass system used by commuters in the northeastern United States conducted such a trial at two McDonald’s drive-thru windows on Long Island, NY. In the American Banker, Wade (2004) reported, however, that the co-owner of these fast food restaurants had no plans to accept E-ZPass at the four other McDonald’s he owned; that he was more excited about the possibility of accepting credit cards because they are more ubiquitous; that the E-ZPass system was more complicated to set up; and that it would be an uphill challenge to reach broad-based adoption. Part of the difficulty relates to the technology. The transponder that drivers use to store value and record tolls is about the size of a deck of cards, which makes it practical for tolls and drive-thru restaurants but not for a typical retail environment. Moreover, a widespread number of merchants would
have to bear the cost of installing smart card readers and a critical mass of consumers would have to prefer the device over other payment alternatives to make it successful.

While we have seen that some programs require smart card usage, this strategy does not guarantee success. The Ohio Direction Card program, which required EBT recipients to access benefits through the Direction Card, had the highest average daily transaction values of the six e-purse programs surveyed, well above the one million users that Goldfinger (1998) estimated would be necessary to offset infrastructure costs. However, Ohio’s EBT system failed on account of the higher cost of the offline infrastructure relative to online systems. The Mcard program failed for the same reason. Therefore, a second key factor in a successful smart card program is overall cost compared with other payment alternatives. As such, the Navy Cash™, EagleCash, and UCF Card programs should be monitored over time. If the overall costs of these programs are higher than other payment technologies, their long-term success may be in jeopardy. This concern may be mitigated to some extent for Navy Cash™ because of the security that the card provides for ship check in.

A third key ingredient for a successful e-purse program is compelling incentives for consumers to use and for merchants to accept the new payment device frequently. In terms of Chakravorti’s (2004) framework, Octopus offers consumers and merchants simultaneous benefits that are not provided by existing payment systems. For consumers, Octopus completes transactions faster than cash, provides an automatic reload feature, offers loyalty programs, and is the only e-purse surveyed that allows the purchase of goods up to a negative value. Merchants that accept the Octopus card benefit from quicker transaction times and ad hoc loyalty programs, as well as lower fees compared with those for credit cards. Nevertheless, a handful of retailers have found that the merchant transaction fees are too high and have withdrawn from the program.

When analysts predicted the success of e-purses during the 1990s, the main benefit envisioned was a cash substitute for small-value purchases. Octopus concentrates on the micropayments environment and provides merchants with reduced cash handling costs. However, Van Hove (2001) argues that merchants experience increased costs in the short term by supporting two separate infrastructures until a critical mass of e-purse users is established. The Navy Cash™ program was successful in substituting cash entirely for crew members on board ships, resulting in significant cost savings. Despite the University of Michigan’s desire to replace coin-intensive transactions like those for pay phones, parking meters, and parking facilities, Mcard’s technology was unable to support these devices.

In the United States, credit cards rather than e-purses may fill the micropayments gap. For example, some parking meters in Chicago accept credit cards, and the credit card industry plans to continue introducing contactless cards to speed payment. JPMorgan Chase & Co., the largest credit card issuer in the United States, has announced that it will issue millions of new contactless credit cards. American Express Co. plans to issue its contactless card, ExpressPay, to new customers in June 2006. More importantly, Visa, MasterCard, and American Express have agreed to a standard that enables properly equipped POS readers to accept contactless credit cards (Sidel, 2005). The success of contactless credit cards in filling the micropayments gap relies on two factors: a critical mass of consumers that are able to find enough chip reading terminals to use these cards frequently and a widespread number of retailers that view fees for small-value transactions on these cards affordable.

The final critical ingredient for a successful smart card program is technology that is well tested and addresses standards issues before rollout. As seen in the Mcard program (see appendix), technological difficulties decreased the adoption rate and contributed to the overall failure of the program. In contrast, Octopus had very few problems in the initial rollout, and after 7-Eleven’s implementation in the retail sector, Hong Kong’s consumers pressed other merchants to accept the card as well. The UCF Card program reflects a similar desire from merchants and users to expand card usage; however, the UCF Card program is the only one with multiple e-purses and requires merchants to select the correct e-purse from which to deduct a transaction. A more streamlined technology that includes a single e-purse may increase adoption, but at the same time, a single e-purse may negatively affect parents’ sense of control over funds being spent for school-related purposes.

Octopus also established interoperability between the various transportation providers before the program was launched, which smoothed its expansion into the retail sector. Still, Octopus may face interoperability issues in the future as Asia migrates to Europay/MasterCard/Visa (EMV) compliant credit cards. EMV facilitates the introduction of chip technology into the international payment systems by developing joint specifications for issuance, acceptance, and interoperability of chip-based debit and credit card transactions. An ePaynews.com article (Trintech Group Plc,
2004) reports that the Chinese banking industry is scheduled to complete a Chinese version of the EMV standard prior to issuing EMV chip cards.

It is difficult to anticipate what impact the EMV initiative might have on Octopus. One possibility is that merchants may be reluctant to have two separate smart card readers—one for Octopus cards and one for EMV credit and debit cards. Instead, they may want the computer chips on Octopus cards to be EMV compliant and POS terminals for the Octopus card and for EMV credit and debit cards to be integrated. This option will likely result in additional costs related to issuing new cards and to integrating existing card readers for both Octopus Cards Ltd. and its merchants. Some of these costs will likely be passed on to consumers. Another possibility is that the relatively low cost of Octopus transactions for merchants may offset the expense of maintaining dual infrastructures.

Military cards may also face interoperability and cost issues because of some new initiatives in the United States, such as the Government Smart Card (GSC) initiative and Homeland Security Presidential Directive (HSPD) 12. GSC is designed to adopt smart card technology for every federal employee for a wide range of purposes, such as building access. HSPD 12 establishes government-wide standards for secure and reliable forms of identification issued by the federal government to its employees and contractors (White House, Office of the Press Secretary, 2004). Open standards on how to build a smart card infrastructure across all federal agencies are being published, and presumably, military cards will need to conform to these standards.

Conclusion

Of the six e-purse programs reviewed, Octopus and Navy Cash™ have come closest to creating the cashless world foretold by many analysts in the 1990s. While the Navy Cash™ program entirely replaces cash on board ships equipped with the system, the scale of the program is far smaller than that of Octopus, which used the transportation industry to achieve critical mass. The Octopus model may be difficult to replicate here because the U.S. population is much larger, is more geographically dispersed, and does not rely as heavily on public transportation. In addition, the United States is a less cash-intensive society compared with Hong Kong, where an e-purse for small-value transactions may have more utility. The United States also has an efficient, advanced, and inexpensive telecommunications infrastructure for debit, credit, payroll, and gift cards that has relatively low levels of fraud.

One way to establish critical mass is to require use of smart cards over other payment alternatives, as seen in the Direction Card program in Ohio. However, even if usage is mandatory, the overall cost of a smart card system relative to other payment options is critically important to its success. Required usage is also impractical in an open-loop environment where consumers are accustomed to payment choices.

Successful e-purse programs also provide consumers and merchants with powerful incentives to use and accept the card. Octopus has been highly successful in this regard; still, a handful of merchants have found the fees too high and defected from the program.

Finally, technology that is well tested and addresses standards issues before rollout is another factor in a successful e-purse program. Inferior technology helped doom the MCard program, while Octopus provides a highly interoperable system with a low error rate.
## APPENDIX: E-PURSE PROGRAMS

<table>
<thead>
<tr>
<th>E-purse</th>
<th>Octopus card</th>
<th>Direction Card</th>
<th>Mcard</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Status</strong></td>
<td>Successful</td>
<td>Failed</td>
<td>Failed</td>
</tr>
<tr>
<td><strong>Cards issued</strong></td>
<td>12,700,000</td>
<td>3,203,066</td>
<td>92,000</td>
</tr>
<tr>
<td><strong>Average usage per transaction</strong></td>
<td>$1.00</td>
<td>$29.50</td>
<td>$0.43</td>
</tr>
<tr>
<td><strong>Average daily transaction value</strong></td>
<td>$1,411,000$^a</td>
<td>$2,765,102$^b</td>
<td>n.a.</td>
</tr>
<tr>
<td><strong>Locations accepting card</strong></td>
<td>Over 50,000</td>
<td>5,123</td>
<td>342</td>
</tr>
<tr>
<td><strong>Reader</strong></td>
<td>Contactless</td>
<td>Contact</td>
<td>Contact</td>
</tr>
<tr>
<td><strong>How it works</strong></td>
<td>Choice of three cards: On-Loan Anonymous costs HK$150 ($19) and includes a HK$50 deposit and a HK$100 stored value amount; Sold Octopus costs HK$70 and has no stored value; and Personalized Octopus costs HK$100 and includes a HK$50 deposit, a HK$20 printing and handling charge, and a stored value amount of HK$30. Consumer can automatically reload personalized cards through credit card or bank account, and most banks and credit card issuing companies offer bonus points or cash rebate for automatic reload feature. Consumer can reload all three cards with cash.</td>
<td>Ohio EBT recipients receive cards for free and are required to load their benefits at any three stores of their choosing or at the local food stamp office.</td>
<td>Card was required ID on campus and included microchip for payments, bar code for checking out library books, and magnetic stripe for building access. The magnetic stripe also functioned as a closed-loop debit card, called BankStripe, linked to a First of America checking account. Card also had calling card capabilities. Value could be added to the e-purse at 23 CashChip machines on campus.</td>
</tr>
<tr>
<td><strong>Maximum stored value</strong></td>
<td>HK$1,000 ($128)</td>
<td>n.a.</td>
<td>$50</td>
</tr>
<tr>
<td><strong>Merchant incentives</strong></td>
<td>Reduced cash handling, quicker transaction times, marketing information, lower fees than credit cards.</td>
<td>Free terminal installation, training, maintenance. No transaction costs.</td>
<td>Lowered cost of terminal and merchant discount fees from 4 percent to 2.8 percent.</td>
</tr>
<tr>
<td><strong>Consumer incentives</strong></td>
<td>Speed, convenience, purchases up to negative value, security, loyalty discounts.</td>
<td>High levels of customer service, standardized procedures.</td>
<td>Coupons, free transfers from checking to chip, discounts when used instead of cash.</td>
</tr>
<tr>
<td><strong>Other factors related to success or failure</strong></td>
<td>Cooperation of major transportation companies, leverage of transportation industry to establish critical mass, reliable and interoperable technology.</td>
<td>Overall costs, interoperability.</td>
<td>Malfunctioning technology, long transaction times, poor training, low merchant discounts, costly readers.</td>
</tr>
<tr>
<td><strong>Adoption</strong></td>
<td>Voluntary</td>
<td>Mandatory</td>
<td>Voluntary</td>
</tr>
<tr>
<td><strong>Other payment instruments accepted at point of sale</strong></td>
<td>Primarily cash, but credit cards, debit cards, checks at some locations.</td>
<td>Cash, checks, credit cards, debit cards.</td>
<td>Cash, checks, credit cards, debit cards.</td>
</tr>
<tr>
<td>E-purse</td>
<td>UCF Card</td>
<td>Navy Cash™</td>
<td>EagleCash</td>
</tr>
<tr>
<td>-------------------</td>
<td>----------------</td>
<td>-----------------</td>
<td>---------------</td>
</tr>
<tr>
<td><strong>Status</strong></td>
<td>Successful</td>
<td>Successful</td>
<td>Successful</td>
</tr>
<tr>
<td><strong>Cards issued</strong></td>
<td>120,000</td>
<td>68,883&lt;sup&gt;1&lt;/sup&gt;</td>
<td>56,659</td>
</tr>
<tr>
<td><strong>Average usage per transaction</strong></td>
<td>$3.98</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td><strong>Average daily transaction value</strong></td>
<td>$6,691</td>
<td>$25,000</td>
<td>$15,000</td>
</tr>
<tr>
<td><strong>Locations accepting card</strong></td>
<td>359</td>
<td>892,000 ATMs and 32 million merchants worldwide.</td>
<td>n.a.</td>
</tr>
<tr>
<td><strong>Reader</strong></td>
<td>Contact</td>
<td>Contact</td>
<td>Contact</td>
</tr>
<tr>
<td><strong>How it works</strong></td>
<td>Students and parents add value through the WebRevalue online service, fax, mail, or phone. Students and parents can also visit the UCF Card Services office and provide cash, check, debit cards, or credit cards. Students and parents can load money to the first e-purse by using Cash-to-Chip machines on campus. Transfers among e-purses require a visit to the UCF Card Services office.</td>
<td>Onboard purchases and chapel donations are made through the chip. Purchases ashore, over the phone, or on the Internet are made through the magnetic stripe at any location accepting MasterCard. Onboard cashless ATMs can be accessed to check balances; move money from stripe to chip and vice versa; access bank or credit union accounts to transfer money; complete a ship “check in” when reporting aboard for permanent or temporary duty; and transfer money from one crew member’s account to another.</td>
<td>Value added through payroll deposits, bank account transfers, and checks. Departing staff turn in card, and any remaining amount is refunded in cash. Cards have an expiration date, so staff who forget to turn in their card have the remaining balance deposited directly to their bank accounts.</td>
</tr>
<tr>
<td><strong>Maximum stored value</strong></td>
<td>E-purse 1: $100, for on-campus purchases at vending machines, photocopiers, laundries, and printing facilities. E-purse 2: $5,000, for tuition and on-campus purchases at the book store; PIN protected. E-purse 3: $5,000, for retail purchases on and off campus; PIN protected.</td>
<td>$1,000; PIN protected, although purchases under $25 from vending machines do not require PIN.</td>
<td>$9,999.00</td>
</tr>
<tr>
<td><strong>Merchant incentives</strong></td>
<td>Access to nearby student population.</td>
<td>None</td>
<td>None</td>
</tr>
</tbody>
</table>
### E-PURSE PROGRAMS (CONTINUED)

<table>
<thead>
<tr>
<th>E-purse</th>
<th>UCF Card</th>
<th>Navy Cash™</th>
<th>EagleCash</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Consumer incentives</strong></td>
<td>Security, discount at vending and copy machines, off-campus restaurant options.</td>
<td>Free card, no transaction fees, greater security, access to home bank accounts.</td>
<td>Free card, payment instrument for use in hostile territories with weak banking and telecommunications infrastructures, speed, security.</td>
</tr>
<tr>
<td><strong>Other factors related to success or failure</strong></td>
<td>Gives parents some control over funds being spent for school-related purchases.</td>
<td>Reduces cash handling, fiduciary, and reporting costs.</td>
<td>Decreases costs of transporting and safekeeping currency and reduces the risk of U.S. currency falling into the hands of terrorists.</td>
</tr>
<tr>
<td><strong>Adoption</strong></td>
<td>Mandatory for computer lab; voluntary for other uses.</td>
<td>Voluntary</td>
<td>Voluntary</td>
</tr>
<tr>
<td><strong>Other payment instruments accepted at point of sale</strong></td>
<td>Some locations cash only; some locations cash, checks, credit cards, debit cards.</td>
<td>None for crew members; visitors may be allowed to use cash and checks.</td>
<td>Cash, checks, credit cards, debit cards.¹</td>
</tr>
</tbody>
</table>

¹The U.S. Army Finance Command has issued a communication that EagleCash is the standard in the field, but the decision to accept other payment instruments at the point of sale rests with local command.

Notes: Amounts in U.S. dollars, unless stated otherwise. n.a. means not available.
Smart cards are credit-card-sized plastic cards embedded with a microchip that are more difficult to counterfeit and can store more information than magnetic stripe cards.

The geographic concentration of the potential users of an e-purse is also important. For example, it would be cheaper to support one million smart card users in a concentrated market like Hong Kong versus one million smart card users in the whole of China.

Hong Kong has had a stable exchange rate of HK$7.8 to US$1 since the mid-1980s; its GDP was US$158 billion in 2003.

For a detailed discussion of the New York City trial, see Van Hove (2001).

While some retailers and credit card companies, such as Target Corporation and American Express, have added computer chips that do not have stored-value capabilities to their cards, their customer programs have either failed or have not established a critical mass of users.

For a discussion of the open-loop Mondex trial in Hong Kong, see Westland et al. (1998). Electronic Transaction Association (2001) discusses the Atlanta Olympics e-purse trial. Again, for a detailed discussion of the New York City trial, see Van Hove (2001).

Total Octopus transactions in 1998 are unavailable. Using current statistics of 8.3 million daily Octopus transactions, this would represent a failure rate of 0.01 percent.

There have been no announcements of Wyoming’s intention to discontinue its smart card program, which is used to administer the EBT program and the Women, Infants, and Children (WIC) program. An industry source indicates that if an online WIC program pilot by the state of Washington proves to be a success and a good business strategy, then Wyoming will be open to the best business case when it rebids its EBT services.

Out-of-state stores accepting smart cards are generally located near Ohio’s and Wyoming’s state borders.

The Mondex’s concurrent programs included Entrée Plus, BankStripe, a calling card program, and a chip program. Entrée Plus transactions were deducted from the school’s pre-established funds pool. BankStripe transactions were deducted from the cardholder’s First of America checking account. Calling card transactions were billed in one of two ways—on the student’s phone bill or on a prepaid basis through Ameritech. Chip transactions were deducted from the stored value on the card (Doyle, 2005).

Enhancements to the Mcard were offered late in the program, but failed to have a positive effect on usage. These included migrating to an open debit card network, offering free funds transfers from First of America checking accounts to the chip, providing visitors with loaded Mcards for meals and other sundries, promoting merchants in monthly newsletters, and reducing the cost of terminals. Other incentives included distributing coupons from participating merchants to students paying with the Mcard, offering students who watched a training video about the card $5 on the chip, and providing vending machine and merchant discounts when the card was used instead of cash (Doyle, 2005).

Unless indicated otherwise, information on the UCF Card is from the University of Central Florida (2005).

Additional information was taken from the U.S. Department of the Treasury, Financial Management Service (2005).

A third program, EZpay, provides all U.S. Army, U.S. Air Force, and U.S. Marine basic trainees with a disposable EZpay card with a fixed amount of electronic currency for purchases at base stores, beauty and barber shops, cleaners, gift shops, museums, phone centers, photo stores, and video stores. The amount loaded on the card is deducted from the trainee’s first paycheck (Mackenzie, 2004).

After paying for tolls in advance, drivers are issued a transponder about the size of a deck of cards that is mounted on the windshield and allows them to pass through a tollbooth without stopping the vehicle. The card can be linked to a consumer’s debit or credit card to automatically replenish funds.

Malfunctioning technology also impeded card adoption in the Mondex and Visa Cash trial in New York City.
REFERENCES


Balaban, Dan, 2005, “Road to retail transactions is a long one for transit card operators,” CardTechnology, March 1.


Cheng, Cindy, 2004, interview with sales and marketing director, Octopus Cards Ltd., by Carol L. Clark, May 12.


Doyle, David, 2005, interview with marketing and sales coordinator, Mcard Administration, Treasurer’s Office, University of Michigan—Ann Arbor, by Carol L. Clark, February 15.


Kidder, Tamara, 2005, interview with UCF card program manager by Carol L. Clark, March 2.


Scaggs, John, 2005, interview with state of Ohio EBT director by Carol L. Clark, January 20.


Straw, Barbara C., 2005a, interview with program director, Navy Cash, U.S. Navy, by Carol L. Clark, September 16.

———, 2005b, interview with program director, Navy Cash, U.S. Navy, by Carol L. Clark, April 15.

Tai, Eric, 2005, interview with chief executive officer, Octopus Cards Ltd., by Carol L. Clark, January 31.


Wong, Joey, 2005, interview with public relations manager, Octopus Cards Ltd., by Carol L. Clark, September 21.

Introduction and summary

Macroeconomics explains the behavior of broad measures of economic activity, such as the total amount of income produced in the economy or the overall average of the prices households pay for the goods and services that they consume. Such aggregates reflect the confluence of many different decisions in product and financial markets. Households decide how much to work and what to spend; businesses commit to capital investment projects; foreigners decide on their demand for U.S. products; and governments determine taxes and spending.

To discern how particular events influence macroeconomic performance, researchers and policymakers sometimes turn to large-scale econometric models. These are collections of statistically estimated equations, which describe a wide range of economic decisions. The results from these equations are then aggregated into macroeconomic outcomes. Large-scale models can address a wide range of relevant questions, but their size comes with a cost. To make estimation and aggregation practical, they require numerous restrictions. Sims (1980) argued that these restrictions often lack theoretical justification, which can cast doubt on some of their predictions. Furthermore, the complexity of large models sometimes makes it difficult to determine the relevant factors underpinning their results.

An alternative approach uses the behavior of broad macroeconomic aggregates to identify a few fundamental sources of change in aggregate production and prices. We follow this methodology and build a small-scale econometric model designed to separate the influence of permanent and transitory factors on the level of economic activity. This distinction can usefully inform policymakers’ decisions. For example, in Taylor’s (1993) formulation of monetary policy rules, the interest rate responds to a transitory deviation in output from its long-run potential level but not to simultaneous movements in both actual and permanent output.

Because the model is small and makes few assumptions, it cannot identify a large number of independent factors that may affect economic outcomes. However, its simplicity is also a virtue—the model’s results have simple interpretations and are robust to a number of specification issues that can plague large-scale systems.

Even small econometric models like ours require assumptions to identify the sources of economic fluctuations. For this purpose, we employ Friedman’s (1957) permanent income theory of consumption. The central prediction of this theory is that forward-looking households consume a constant fraction of their permanent income—the sum of their financial wealth and the financial value of their current and future labor earnings. This leads us to use the identifying assumption that all permanent changes in nondurable consumption reflect changes in households’ permanent income.

The model links nondurable consumption with other economic expenditures using the observation of balanced growth: households’ expenditures on very broad categories of goods and services do not drift apart over time. With balanced growth, permanent changes in the consumption of nondurable goods and services eventually lead to equally sized changes in other spending. The model we estimate imposes balanced growth across consumption of nondurable goods and services, durable goods expenditures, and total private savings.4

We estimate our model using data on the U.S. economy from 1983:Q1 through 2005:Q2. This sample deliberately omits the period of high and variable inflation beginning in the early 1970s, so the model’s...
dynamics reflect the monetary policy from the period of low and stable inflation that followed Paul Volcker’s chairmanship of the Federal Reserve. Furthermore, since the early 1980s the volatility of many macroeconomic aggregates has been much lower than it was in the immediate post-Korean War era. Our model will reflect the structure of the economy during only this less volatile period. Of course, using such a short sample limits the number of parameters that we can estimate. This also leads us to highly value parsimony.

The remainder of this article proceeds as follows. We next review the economic theory that guides our forecasting model’s specification. Then we examine the implications of this theory for our empirical model and describe the data. After this, we present the estimated model and its forecasts. Finally, we discuss the use of our model to measure and forecast the gap between actual output and its long-run potential level.

**Theoretical foundations**

Theoretically oriented macroeconomics guides the specification of our forecasting model, so we begin our analysis with a review of two of its important results. First, consumption of nondurable goods and services should be helpful for forecasting other macroeconomic quantities. Second, the shares of households’ expenditures on broad classes of items do not exhibit permanent changes.

We derive both of these results from the theory of a forward-looking household’s optimal consumption and savings decisions, developed originally by Friedman (1957). The household values two goods, consumption of nondurable goods and services, \( C_t \), and the service flow from a stock of durable goods, \( S_t \). For simplicity, we assume the service flow from durable goods is equal to the stock.\(^3\)

The household’s resources are its labor income and its initial wealth. Labor income in period \( t \) is \( Y_t \), and this is potentially random. The household places all of its financial wealth in a single asset, risk-free bonds. The face value of the bonds at the beginning of period \( t \) is \( B_0 \), and the household can purchase bonds that come due in period \( t + 1 \) at the price \( 1/R_t \). We assume that each period is a calendar quarter, so the implied annual interest rate on these bonds in percentage points is \( 400 \times (R_t - 1) \). Both bonds and labor income are denominated in units of the nondurable good. The household can trade these resources for durable goods at the relative price \( P_t \). Let \( X_t \) and \( \delta \) represent the household’s purchases of durable goods and the constant rate at which they depreciate so that \( S_t = (1 - \delta) S_{t-1} + X_t \).

Given the household’s initial financial wealth \( B_0 \) and the value of its used durable goods stock \( (P_0(1 - \delta) S_0) \), the household allocates its resources across the consumption of nondurable goods and services, durable goods purchases, and bonds in order to maximize the expected utility function,

\[
E_0 \left[ \sum_{t=0}^{\infty} \beta^t \ln \left( C_t S_t^{\alpha} \right) \right],
\]

subject to the sequence of budget constraints,

1) \( C_t + P_t X_t + B_{t+1} / R_t = Y_t + B_t \).

Here, \( E_0[Z] \) denotes the mathematical expectation of the random quantity \( Z_t \), calculated using information available at time 0.\(^3\) Increasing \( \beta \) increases the household’s value of future consumption relative to current consumption. In this sense, \( \beta \) measures the household’s patience. The exponent \( \theta \) lies between 0 and 1: This ensures that utility increases and marginal utility falls with increasing consumption of either \( C_t \), or \( S_t \). The larger \( \theta \) is, the more the household values nondurable goods relative to durable goods.

The role of expected future income in current consumption decisions can be most easily appreciated by replacing the sequence of budget constraints (equation 1) with a single unified budget constraint,

2) \( \sum_{t=0}^{\infty} \hat{R}^{-t} E_0 \left[ C_t + P_t X_t \right] = B_0 + \sum_{t=0}^{\infty} \hat{R}^{-t} E_0 \left[ Y_t \right] \).

Here, \( \hat{R} \) is the interest rate on a \( t \)-period risk-free bond sold in period 0. The left-hand side equals the present value of the household’s expenditures on nondurable and durable consumption goods, while the right-hand side equals the present value of its assets—its financial wealth plus the value of its current and future labor income.\(^4\) The feature of equation 2 that is important for our purposes is that an increase in \( E_0[Y_t] \) increases the resources available for consumption at all dates, including the present.

**Consumption as a predictor of future income**

To demonstrate the forecasting power of nondurable consumption for income, we begin by characterizing the household’s savings decision. If \( C_t \) and \( B_{t+1} \) are optimal, then the household cannot make itself better off by reducing \( C_t \) slightly, using the foregone consumption to purchase bonds and consuming the principal and interest in period \( t + 1 \). The utility cost of slightly decreasing \( C_t \) is \( \theta / C_t \), and the future utility benefit from temporarily increasing savings is \( R \theta / C_{t+1} \). This benefit is potentially random, because the household’s actual choice of \( C_{t+1} \) could be affected by random
changes in labor income or interest rates in period $t + 1$. Hence, we wish to compare the cost of foregone consumption with the expected future benefit, $E[\frac{R_t}{C_{t+1}}]$. Discounting this back to period $t$ and equating it to the current-period utility cost yields

$$3) \frac{1}{C_t} = \beta R_t E_t \left[ \frac{1}{C_{t+1}} \right].$$

Because the consumer’s choices of $C_{t+j}$ and $B_{t+j}$ must also be optimal, given the information available at time $t + j$, a version of equation 3 must hold in all future periods. That is,

$$4) E_t \left[ \frac{1}{C_{t+j}} \right] = \beta E_t \left[ R_{t+j} \frac{1}{C_{t+j}} \right].$$

Equation 4 arises from taking expectations of both sides of equation 3 after first increasing the dates of all variables by $j$ periods.

Equation 3 implies that changes in the household’s expectations of future income directly influence $C_t$. To see this, hold the interest rate at some constant value and suppose that $E_t[Y]$ increases so that the right-hand side of equation 2 rises by 1 percent. If the household spends all of this additional income on nondurable goods, then the only solution to the optimality conditions, equation 3 and equation 4, is to increase $C_t$ and all of its future values by 1 percent. Allowing the household to use some of this extra income to purchase durable goods changes the magnitude of the consumption response, but not the result that the current and future percentage responses are the same. In this sense, $C_t$ is a forward-looking variable that should be informative about the household’s expectations.

A second implication of equation 3 is that no macroeconomic variable can improve a forecast of $C_{t+1}$ that already uses $C_t$. This can be seen by multiplying both sides of equation 3 by $C_t$ setting $R_t = \beta^{-1}$, and rearranging terms to get

$$5) E_t \left[ \frac{C_{t+1}}{C_t} - 1 \right] = 0.$$ 

Equation 5 embodies Hall’s (1978) result that no information available at time $t$ is useful for forecasting the growth rate of nondurable consumption. If we relax the strong assumption that $R_t = \beta^{-1}$, then the appropriately modified version of equation 5 implies that the interest rate is the only variable with information about the growth rate of consumption. Together, these results characterize the role of $C_t$ in macroeconomic forecasting: Nondurable consumption is informative about future income, but only the interest rate can help predict its growth.

**Durable goods**

To tie the evolution of nondurable consumption with other macroeconomic aggregates, we begin by characterizing the household’s utility-maximizing durable goods purchases. Again, consider a small change to the household’s optimal expenditures. Suppose that the household sells $P_t$ units of nondurable consumption for one unit of durable consumption, holds that durable good until the next period, and then sells it on a used durable goods market. Since the household’s initial choices were optimal, this small adjustment cannot increase utility. The cost of this adjustment is $P_t \beta / C_t$, the utility value of the foregone nondurable consumption. There are two benefits. First, the consumer enjoys the utility of the service flow from the additional durable goods, $(1 - \delta) P_t$. Second, next quarter the consumer can expect to sell the depreciated durable goods for $(1 - \delta) P_{t+1}$ units of nondurable consumption goods which are worth marginal utility of $\beta \delta / C_{t+1}$. Equating the costs with the expected benefits, assuming that $P_t$ grows at the constant rate $\gamma$, and rearranging terms yields

$$6) \frac{P_t S_t}{C_t} = \frac{1 - \delta}{\delta} \left( \frac{R_t - (1 - \delta) \gamma}{R_t} \right)^{-1}.$$ 

On the right-hand side of equation 6, the term $(R_t - (1 - \delta) \gamma)/R_t$ can be interpreted as the upfront cost of renting $1/P_t$ units of durable goods. This equals the period $t + 1$ expense of repaying the interest and principal from borrowing the purchase price minus the resale value discounted back to period $t$. If the interest rate, preferences for durable goods (the parameters $\beta$ and $\delta$), and technology ($\delta$ and $\gamma$) remain unchanged, then this rental cost is constant. In this case, the right-hand side of equation 6 is constant, so the ratio of nominal expenditures on nondurable consumption to the value of the durable goods stock also does not change over time. This is one aspect of balanced growth, which we now proceed to examine in more detail.

**Balanced growth**

We call economic growth balanced if it leaves households’ expenditure shares on broad classes of items unchanged in the long run. Balanced growth ties the long-run levels of macroeconomic quantities together, thereby aiding forecasting.
From the end of World War II until the early 1980s, households’ financial savings, their purchases of durable goods, and their expenditures on nondurable goods and services were balanced. This motivated the builders of early general equilibrium business cycle models, such as King, Plosser, and Rebelo (1988), to assume that as income rises, households will continue to consume the same shares of the various goods and services available in the economy.

We present here the theoretical foundations of balanced growth. Suppose that the household’s labor income grows at the constant rate \( Y_{t+1}/Y_t = \mu \), the interest rate equals the constant \( R = \beta^{-1} \mu \), and the price of durable goods equals \( P \) always. We wish to find the utility maximizing choices of \( C_t, S_t \) and \( B_{t+1}/B_t = \mu \). It turns out that the unique choice of \( C_t \) that is consistent with this sequence satisfying \( B_{t+1}/B_t = \mu \) is

\[
C_t = \kappa Y_t/(1-\beta) + B_0 + P(1-\delta)S_{t-1}.
\]

Here, \( \kappa \) replaces a complicated expression of \( \beta \), \( \theta \), \( \mu \), and \( \delta \). The sum multiplying \( \kappa \) is the value at time 0 of the household’s permanent income. This is the sum of its tangible wealth and the present value of its current and future labor income, \( Y_t/(1-\beta) \).

We are also interested in the growth rates of durable consumption expenditures and household savings. Durable goods purchases directly inherit the growth rate of the durable goods stock, because

\[
X_{t+1}/X_t = S_{t+1}/S_t \frac{(S_{t+1}-(1-\delta)S_t)}{(S_t-(1-\delta)S_{t-1})} = \mu \frac{1-(1-\delta)/\mu}{1-(1-\delta)/\mu} = \mu.
\]

Net private savings equals \( Y_t + (R-1)B_t - C_t - PX_t \). All four terms grow at the rate \( \mu \), so their sum does as well. Thus, the model predicts that the household’s expenditures on nondurable goods, its durable goods purchases, and its savings all grow at a common rate, \( \mu \).

**Permanent income shocks**

We interpret the balanced growth path as a description of the household’s long-run choices in the absence of business cycles. However, the solution in the previous subsection can be used to develop intuition about short-run responses to changes in permanent income. To do so, suppose that the household is originally on a balanced growth path so that the solution in equation 7 sets \( C_0/C_{t-1} = \mu \). We say that the household receives a permanent income shock if \( Y_t \) unexpectedly rises above \( \mu Y_{t-1} \), but all other growth rates of \( Y_t \) remain unchanged. Equation 6 requires both \( C_t \) and \( S_t \) to increase by a common percentage. Typically this requires the household to spend some of its savings. Thereafter, \( C_t, S_t \), and \( B_t \) all continue to grow at the rate \( \mu \).

Consider the observable implications of this response for consumption expenditures. The behavior of \( C_t \) following a permanent income shock is consistent with the analysis in the previous subsection on consumption as a predictor of future income: The shock to \( Y_t \) causes a one-time shift up in the path for \( C_t \) and induces no forecastable changes in its subsequent growth rates. Next, consider the shock’s impact on \( X_t \). Purchases in period \( t = 0 \) must cover both depreciation on the existing stock and the increase in the desired stock to \( S_0 \); that is, \( X_0 = S_0 - (\mu-1)S_1 + \delta S_{-1} \). Rearranging terms in this expression allows us to relate the growth rate of \( S_0 \) with that of \( X_0 \):

\[
\frac{X_0 - X_{-1}}{X_{-1}} = \left( \frac{S_0 - S_{-1}}{S_{-1}} \right)/(\mu-1+\delta).
\]

For empirically relevant choices of \( \mu \) and \( \delta \), the term \( \mu-1+\delta \) is considerably less than one. Thus, the given change in \( S_0 \) translates into a much larger percentage change in \( X_0 \). This change is mostly temporary so that \( X_t = (\mu+\delta)S_t \). Hence, the model leads us to expect expenditures on durable goods to be much more volatile than expenditures on nondurable goods. This volatility reflects transitory aspects to growth that arise from the role of durable goods expenditures as an investment in household capital.

**The forecasting model**

The remainder of this article describes the construction and use of a vector autoregression (VAR) forecasting model of the U.S. economy. The three key aggregates from the U.S. National Income and Product Accounts (NIPA) that it includes are privately demanded gross domestic product (GDP), personal consumption expenditures on nondurable goods and services, and personal consumption expenditures on household durables. The model builds on the theory of household consumption and savings in two ways. First, it uses consumption expenditures on nondurables and services to identify shocks to permanent income. Second, it constrains the evolution of durable consumption expenditures and private savings so that the model’s forecasts satisfy balanced growth restrictions in the long run.
The theory identifies the interest rate of a consumption-denominated bond as informative for both consumption growth and durable goods purchases, so the forecasting model also uses information on the real cost of borrowing. The interest rates in U.S. financial markets are typically denominated in dollars, so they must be adjusted by market participants’ expectations of inflation to produce a real interest rate directly relevant for households’ decisions. Instead of explicitly measuring inflation expectations and using these to directly construct the real interest rate, we include the interest rate on federal funds and the inflation rate in the forecasting equations of all variables. If market participants forecast inflation with some linear combination of current and lagged variables in the model, then we can interpret the estimated equations as including market participants’ inflation forecasts. Inflation and interest rates are themselves of substantial macroeconomic interest, so the model includes forecasting equations for them as well.

**Data**

We must first address an important measurement issue before taking our model to the data. The distinction between durable goods and services is subtle. A system of national income accounts can either treat a durable goods purchase as an investment (recorded as an expenditure by the business sector) that yields a flow of rental services to households or as a direct expenditure by the household sector. The U.S. NIPA do not treat this issue consistently. The NIPA treat housing as a business transaction: The construction of a home is recorded as residential investment, and the flow of services from the housing stock is recorded as consumption expenditures. In contrast, the purchase of any other new household durable good is counted only as a consumption expenditure: The service flow from the durable does not appear anywhere in the national accounts. For the purposes of this article, we assume that the household sector purchases all durables directly. In practice, this means that we subtract the service flow from housing from the NIPA consumption data and from GDP when constructing our measures of nondurables and services consumption and total private income. We also measure household expenditures on durables as the sum of expenditures on durables and residential investment.

Next, it is helpful to examine the histories of the five variables of interest. We begin with the model’s core variable, consumption of nondurables and services. Figure 1 plots $\Delta C_t = \ln(C_t/C_{t-1})$. For growth rates sufficiently close to zero, this approximately equals $(C_t - C_{t-1})/C_t$. Its mean is 3.25 percent. This average growth rate does not drift over time, which indicates that there are no long-run changes in the growth rate of consumption that could confound our analysis. However, the volatility of $\Delta C_t$ diminished substantially in the 1980s. Its standard deviation over the 1954–82 period was 1.65 percent; since then, the standard deviation has been 1 percent. McConnell and Perez-Quiros (2000) documented that the variances of most NIPA expenditure categories declined at about this time.

To impose long-run balanced growth on the model’s forecasts, we work with the ratios of expenditures on durable goods and nominal privately demanded GDP to nominal nondurable consumption. Figures 2 and 3 plot the logarithms of these ratios from 1954:Q1 through 2005:Q2. According to the balanced growth hypothesis, neither ratio should persistently drift away from its mean. This was the case through the early 1980s, but since then both of these series have persistently declined. One way to measure the drift in these data is shown in the dashed lines, which plot the averages of the series’ previous 40 quarters’ values. These moving averages are stable in the 1960s and 1970s, trend

![Figure 1](image-url)
down in the 1980s, and stabilize to some degree in the 1990s. The shift down in the values of the private income ratio reflects the well-known decline in the U.S. savings rate.

Since nondurable consumption is our fundamental indicator of the income process, we measure inflation using the corresponding implicit price index. Panel A of figure 4 plots this index’s inflation rate. Over the post-Korean War sample, inflation underwent some dramatic changes. It rose in the late 1960s, surged with the oil shocks in the 1970s, and then fell in the 1980s and 1990s. The inflation movements over the past 25 years, however, are small relative to the swings that occurred between the late 1960s and early 1980s. Panel B of figure 4 plots the nominal interest rate on federal funds. Its changes track inflation’s well, but it declined from its peak in the early 1980s somewhat more slowly.

Sample period

The decline of macroeconomic volatility and persistent movements in nominal expenditure ratios pose problems for an econometric model based on stable relationships. Recent research has documented and tried to explain the marked changes in a wide range of many macroeconomic aggregates that apparently have occurred since the early 1980s. Campbell and Hercowitz (2004) argue that the decline in macroeconomic volatility reflects in part greater household access to credit markets. Other structural changes that we expect to influence household decisions in the long run include the elimination of consumer credit interest deductions and the development of secondary markets for mortgage debt. Structural changes in production and distribution may also be a factor; for example, Kahn, McConnell, and Perez-Quiros (2002) argue that improved inventory management technology has helped reduce the volatility of GDP. Monetary policy has also been a source of structural change. The rise in inflation that began in the late 1960s reflected in part economic policies that favored economic growth over inflation stabilization, while Paul Volcker’s increased attention to inflation in 1979 began the transition to the current era of very low inflation.

Explaining structural changes such as these lies outside of the scope of our forecasting model. Nevertheless,
monetary policy regime and factors that may have reduced economic volatility also plausibly changed the way economic agents respond to a variety of economic shocks.

This sample period is much shorter than those used to estimate most macroeconomic forecasting models, which generally start in the late 1960s or even earlier. This greatly limits the number of parameters that we can estimate with sufficient statistical precision. This need for parsimony dictated by the use of a short sample complements the balanced-growth arguments for considering a model that forecasts only a few broad NIPA aggregates. It also leads us to use exclusion tests to further restrict the number of model parameters.

The stochastic simultaneous equations model

The first step in the construction of our forecasts is the estimation of a small structural model of the U.S. economy. The model consists of a system of simultaneous equations for the five variables of interest. Because the equations allow a number of shocks to simultaneously influence multiple variables, we refer to this system as the stochastic simultaneous equations model. This model imposes restrictions from theory in order to identify permanent and transitory shocks to income. In this sense, the model is structural.

The model’s first equation is for the growth rate in the consumption of nondurables and services. This equation also defines the permanent income shock, $\psi_r$:

$$
\Delta c_i = \mu + \sum_{j=1}^{m} \alpha_{i,j} \Delta c_{i,j-1} + \\
\sum_{j=0}^{m} \alpha_{i,0} \pi_{t-j} + \sum_{j=0}^{m} \alpha_{i,1} r_{t-j} + \epsilon_i,
$$

where $\pi_i$ and $r_i$ denote the annualized one-quarter inflation rate and the federal funds rate. This specification embodies the theoretical restriction that only consumption and interest rates predict future consumption. However, the presence of $\pi_i$ and $r_i$ on its right-hand side provides a channel for transitory shocks to immediately impact $\Delta c_i$. 

Their presence does not eliminate the value of imposing balanced growth on our forecasts because after some adjustment period, the nominal ratios will settle down to their new balanced growth path. Accordingly, we decided to make two adjustments to our model. First, we forecast the deviations of the two nominal ratios from their 40-quarter moving averages instead of the nominal ratios themselves. This procedure removes the drift in these ratios’ means arising from structural change. Second, we use only data from 1983:Q1 through the present. We do so because the change in
The model’s long-run restrictions

The restriction that only the permanent income shock can permanently change the level of nondurable consumption implies a set of restrictions on the unknown coefficients in equation 8. We derive these restrictions given the assumption that \( \sum_{i=0}^{m_c} \alpha_c < 1 \). This implies that no shock permanently changes the growth rate of nondurable consumption. This is a prediction of most general-equilibrium business cycle models.

To begin, suppose that a particular transitory shock occurs in quarter \( t \) that impacts both inflation and interest rates. Denote the changes in the level of nondurable consumption, inflation, and the interest rate in quarter \( t \) due to the shock with \( \nabla_c, \nabla \pi, \text{ and } \nabla r \). We assume that the long-run responses of inflation and interest rates to the shock both equal zero so that \( \nabla \pi = \lim_{j \to \infty} \nabla \pi_{rj} = 0 \) and \( \nabla r = \lim_{j \to \infty} \nabla r_{rj} = 0 \). We wish to characterize \( \nabla c = \lim_{j \to \infty} \nabla c_{rj} \) so that we can find the restrictions required to set it equal to zero. From equation 8 we can write \( \nabla c_{rj} \) as

\[
\nabla c_{rj} = \nabla c_{r,j-1} + \sum_{i=1}^{m_c} \alpha_{c} \left( \nabla c_{r,j-1} - \nabla c_{r,j-1} \right) + \sum_{i=0}^{m} \alpha_{z} \nabla \pi_{r,j-1} + \sum_{i=0}^{m} \alpha_{p} \nabla r_{r,j-1}.
\]

Summing these equations for \( j = 0,1,\ldots, M \) yields

\[
\nabla c_{r,M} = \sum_{i=1}^{m_c} \alpha_{c} \left( \sum_{j=0}^{M} \nabla c_{r,j} \right) - \sum_{i=0}^{m} \alpha_{z} \left( \sum_{j=0}^{M} \nabla \pi_{r,j} \right) - \sum_{i=0}^{m} \alpha_{p} \left( \sum_{j=0}^{M} \nabla r_{r,j} \right).
\]

As \( M \) becomes very large, \( \nabla c_{r,\infty} \) and its \( m_1 \) lagged values all approach \( \nabla c \). Because both \( \nabla \pi \) and \( \nabla r \) equal zero, it satisfies

\[
1 - \sum_{i=0}^{m} \alpha_{z} \sum_{j=0}^{M} \nabla \pi_{r,j} = 0 \quad \text{and} \quad 1 - \sum_{i=0}^{m} \alpha_{p} \sum_{j=0}^{M} \nabla r_{r,j} = 0.
\]

Neither \( \sum_{j=0}^{M} \nabla \pi_{r,j} \) nor \( \sum_{j=0}^{M} \nabla r_{r,j} \) necessarily equals zero, so we require both \( \sum_{j=0}^{m} \alpha_{z} \) and \( \sum_{j=0}^{m} \alpha_{p} \) to equal zero for \( \nabla c \) to equal zero. These are the two restrictions we impose on the model.

The equation defines \( \varepsilon \) as the permanent income shock by restricting it to be the only shock that has a long-run impact on \( c \). In practice, \( \varepsilon \) encompasses many fundamentals such as changes in technology, regulation, and access to financial markets. Gathered together, these are all of the factors causing permanent and equalized changes in \( c, X, \) and \( Y \). Other shocks may have transitory effects, but none have any permanent ones.

As a technical matter, as shown in box 1, the two necessary and sufficient conditions for \( \varepsilon \) to be the only shock that has a long-run impact on \( c \) are \( \sum_{j=0}^{m_c} \alpha_{c} = 0 \) and \( \sum_{j=0}^{m} \alpha_{p} = 0 \). These conditions guarantee that any negative influence of a shock to either interest rates or inflation on nondurable consumption growth in the current quarter is offset by a positive influence in the following quarters, so the net effect on \( c \) is zero. We impose these restrictions on our estimates of equation 8’s unknown parameters.

The model’s remaining equations allow \( \varepsilon \) to immediately impact all variables; in particular, \( \varepsilon \) may contemporaneously influence the values of \( \pi \) and \( r \) that enter equation 8. This implies that the equation’s error term and its right-hand side variables are correlated, so we cannot estimate its unknown parameters using ordinary least squares regression. Instead, we assume that there does not exist any information available before time \( t \) that is useful in forecasting \( \varepsilon \). This is a natural restriction to apply to an economic shock. Mathematically, this means that the expectation of \( \varepsilon \) is independent of the lagged values of the data in the model; that is:

\[
E[\varepsilon_j] = E[\varepsilon_{t+j}] = E[\varepsilon_{t-j}] = E[\varepsilon_{t-j}] = 0.
\]

for all \( j \geq 1 \). Here, \( y \) and \( x \) are the deviations from 40-quarter moving averages in the logarithms of the nominal ratios of private GDP and durable goods purchases to nondurable consumption. The conditions in equation 9 suggest estimating \( \mu \) and the \( \gamma \)'s by using the values that set the sample covariances between \( \varepsilon \) and the \( t-j \) dated data equal to zero. However, there are more covariances to set equal to zero.
than there are coefficients to choose. This leads us to use linear combinations of the covariances—which also have an expected value equal to zero—to estimate the equation. We use Hansen’s (1982) generalized method of moments (GMM) estimator, which identifies the particular linear combinations that yield the most precise estimates.\footnote{1}

The model’s second equation is:

\begin{align}
10) \quad \pi_t &= \mu_z + \sum_{j=0}^{m_0} \beta_x \Delta c_{t-j} + \sum_{j=1}^{m_1} \beta_x \pi_{t-j} + \sum_{j=1}^{m_1} \beta_y \tau_{t-j} + \\
&\quad \sum_{j=0}^{m_2} \beta_p y_{t-j} + \sum_{j=0}^{m_2} \beta_p x_{t-j} + \epsilon_{st}.
\end{align}

Because \( \pi_t \) appears on its left-hand side, we call this the inflation equation. The error term in equation 10 has no concrete economic interpretation. Instead, we treat it as a statistical forecast error. Because it is a forecast error, its covariance with any variable that could be used to forecast it equal zero. That is,

\[
E[\epsilon_{st}] = E[\epsilon_{st} \Delta c_{t-j}] = E[\epsilon_{st} \pi_{t-j}] = E[\epsilon_{st} \tau_{t-j}] = E[\epsilon_{st} y_{t-j}] = E[\epsilon_{st} x_{t-j}] = 0
\]

for all \( j \geq 1 \). Thus we can use the corresponding sample covariances to estimates the \( \beta \)'s. In addition, because we restrict \( \epsilon_t \) to be the only shock with a long-run impact on \( c_t \), we also impose the restriction that \( \epsilon_{st} \) is independent of \( \epsilon_t \). Mathematically, this means \( E[\epsilon_{st} \epsilon_t] = 0 \), which is one more covariance to use in estimation. Just as with equation 9, we use the GMM estimator to estimate its unknown parameters.

The specification and estimation of the remaining three forecasting equations proceed similarly. In each of them, \( \tau_{t-j}, y_{t-j}, \) and \( x_{t-j} \) take the role of \( \pi_t \) in equation 10. We selected the lag lengths \( (m_0, ..., m_2) \) for each equation as the smallest value of \( m \) such that the coefficients multiplying an additional lag’s variables are jointly statistically insignificant.

**VAR results**

Table 1 displays the estimated system of simultaneous equations using matrix notation. The system is parsimonious: Only the inflation equation has more than two lags. The coefficient multiplying \( \tau \) in equation 8 and the coefficients multiplying \( \Delta c \) in the inflation and interest rate equations are statistically significant. This implies that transitory shocks can affect consumption growth immediately and that the permanent income shock can affect both inflation and interest rates on impact. However, the coefficients multiplying \( \Delta c \) in the two expenditure-ratio equations are not statistically significant. Thus, the initial changes in nondurable consumption are proportional to the changes in income and durable consumption expenditures.

The apparent ability of transitory economic shocks to influence the growth rate of nondurable consumption sharply contrasts with Hall’s (1978) theoretical prediction that all changes in consumption are permanent. However, Hall’s result depends on the assumption of a constant interest rate. In more general models with a market-determined interest rate, shocks that temporarily change the economy’s productive capability will temporarily affect both the level of consumption and the interest rate. The estimated impact of transitory shocks on consumption growth in our model demonstrates this possibility’s empirical relevance.

Before proceeding to use the model for forecasting, we wish to examine whether it displays evidence of specification error. To do so, we created figure 5, which plots the actual values of the model’s variables along with the estimated model’s forecasts for them given the data known in 1983:Q4.\footnote{12} Structural change in the sample would cause the predicted values to drift away from the actual data. In fact, the data track the initial forecasts well.

**Impulse responses to the permanent income shock**

The permanent income theory makes very specific predictions for the responses of \( \Delta c \) and \( x \) to a permanent income shock. Figure 6 plots the estimated model’s counterparts to these. Each panel plots the response of a variable of interest over time to a single positive permanent income shock equal in magnitude to the shock’s estimated standard deviation. As the theoretical model suggests, consumption of nondurables and services reacts significantly and quickly, it increases about 0.27 percent when the shock hits and achieves its complete increase of 0.44 percent after three quarters.

The adjustment of household capital lasts longer. On impact, \( X_r \) rises about 0.60 percent. The response increases to 1.36 percent after two quarters, remains at this level for about another year, and then slowly moves down toward the same permanent increase found in \( C_r \). This pattern demonstrates how expenditures on durable goods must temporarily rise more than \( C_r \) in order to bring the stock of durable goods into balance with consumption. However, the persistence of the increased expenditures is quite protracted relative to the theoretical model’s predicted transitory increase. We believe that the drawn-out response reflects costs of quickly adjusting the stock of durable goods that are absent from the model.
\[ \begin{align*}
\Delta c_t &= -0.18 \quad -0.10^{**} \quad -1.04^{**} \quad 0.00 \quad 0.00 \\
\pi_t &= 1.29^{***} \quad 0.19 \quad -0.74^{*} \quad 0.34 \quad -0.30 \\
r_t &= 0.54^{***} \quad 0.12^{***} \quad 1.20^{***} \quad 0.31^{***} \quad 0.05^{*} \\
y_t &= 0.16 \quad 0.07 \quad 0.27^{*} \quad 0.91^{***} \quad 0.13^{***} \\
x_t &= 0.71 \quad -0.56^{**} \quad -0.00 \quad -0.06 \quad 0.93^{***} \\
\end{align*} \]

\[ \begin{align*}
\Delta c_{t-1} &= 1.00 \quad -0.08 \quad -0.66^{***} \quad 0.00 \quad 0.00 \\
r_{t-1} &= 2.74^{***} \quad 1.00 \quad 0.00 \quad 0.00 \quad 0.00 \\
y_{t-1} &= 0.56^{***} \quad 0.00 \quad -1.00 \quad 0.00 \quad 0.00 \\
x_{t-1} &= -1.08 \quad 0.00 \quad 0.00 \quad 0.00 \quad 1.00 \\
\end{align*} \]

\[ \begin{align*}
\Delta c_{t-2} &= 0.22 \quad 0.03 \quad 0.38^{***} \quad 0.00 \quad 0.00 \\
r_{t-2} &= 0.85^{**} \quad 0.32^{**} \quad 0.79 \quad 0.13 \quad 0.38^{***} \\
y_{t-2} &= -0.17 \quad 0.04 \quad -0.29^{*} \quad 0.06 \quad -0.25^{***} \quad -0.03 \\
x_{t-2} &= -0.05 \quad -0.01 \quad -0.29^{*} \quad 0.06 \quad -0.16^{***} \\
\end{align*} \]

\[ \begin{align*}
\Delta c_{t-3} &= 0.00 \quad 0.00 \quad 0.00 \quad 0.00 \quad 0.00 \\
r_{t-3} &= -1.40^{***} \quad 0.04 \quad -0.18 \quad 0.07 \quad -0.04 \\
y_{t-3} &= 0.00 \quad 0.00 \quad 0.00 \quad 0.00 \quad 0.00 \\
x_{t-3} &= 0.00 \quad 0.00 \quad 0.00 \quad 0.00 \quad 0.00 \\
\end{align*} \]


\* Significant at the 10 percent level.  
** Significant at the 5 percent level.  
*** Significant at the 1 percent level.

Notes: \( \Delta c_t \) refers to the real logarithmic growth rate of personal consumption expenditures on nondurable goods and nonhousing services; \( \pi_t \) refers to inflation rate measured using that series' chained price deflator; \( r_t \) denotes the nominal interest rate on federal funds, and \( y_t \) and \( x_t \) refer to the logarithmic ratios of nominal private gross domestic product (GDP) and nominal durable goods expenditures to nominal nondurable consumption, adjusted as described in the text. See the text for further notes.

The impulse responses of privately demanded GDP also exhibit a drawn-out cyclical pattern. Initially, \( y_t \) rises by 0.22 percent, it then climbs to 0.41 percent after two quarters — which is very close to its long-run level — and then reverts to 0.25 percent after a steady of remaining at this level for one year. The second and third shocks also exhibit a drawn-out cyclical pattern.
rate rises towards its original level, but remains about 7 basis points below it. In light of the initial deflation, we can interpret the drop in the interest rate as an accommodative policy response to the apparent absence of inflationary pressure.

**Variance decompositions**

How important are permanent income shocks for economic fluctuations? Table 2 reports the fractions of forecast uncertainty due to the permanent income shock and due to the transitory shocks taken together at various forecasting horizons. For consumption, the permanent income shock accounts for a large
percentage of the movement at all forecast horizons; it explains 63 percent of the variance of the one-quarter-ahead forecast error, 73 percent of the four-quarter-ahead error’s variance, and nearly all of the five-year-ahead error. Still, many temporary factors influence the economy, and these explain a substantial fraction of nondurable consumption’s forecast-error variance over the first year or two following the shock.

These results highlight an important difference between our model and Cochrane’s (1994). He identifies the permanent shock by restricting it to be the only factor affecting consumption in the current period. So, by construction, his transitory shock explains
none of the one-quarter-ahead forecast error in consumption. After about ten quarters, however, this shock accounts for about one quarter of the forecast error in consumption, and it remains at this level of importance for all subsequent forecast horizons. In contrast, our identifying assumptions allow transitory factors to influence the near-term forecast, but restrict them from having any influence on the long-run outlook for consumption.

Next, consider the forecast-error variances for the other spending variables in our model. Transitory factors explain most of the forecast errors in durable goods expenditures and total private income over the short and medium terms. For example, even at the two-year horizon, the transitory shocks explain 80 percent of private GDP's forecast-error variance. By construction, the influence of the transitory shocks falls to zero as the forecast horizon increases. However, this takes a long time to occur. At a 20-year horizon, the transitory shocks still account for 36 percent of the forecast error in $Y_t$ and 56 percent of the error in $X_t$.13

The permanent shock has an important influence on inflation and interest rates in the very near term. It accounts for about 40 percent of the one-quarter-ahead forecast error in inflation and nearly 20 percent of that for the interest rate. Its influence on the interest rate drops quickly. At the two-year horizon, it accounts for only 3.7 percent of the forecast-error variance. This is consistent with economic intuition. As we noted previously, the long-run equilibrium real interest rate depends on the trend in the growth rate of consumption but does not change with shocks to the level of output. The influence of permanent income shocks on inflation variance also falls as the forecast horizon increases, but not as rapidly.

**What does the permanent income shock say about history?**

During the 1983–2005 period that we use to estimate our model, the U.S. economy experienced two recessions, a productivity boom, and a decline in inflation. Does the shock to permanent income that we estimate have anything interesting to say about these developments? Consider figure 7. The solid line in each panel plots the deviations of the variable from the forecast based on the data in hand in 1983:Q4 (the forecasts shown in figure 5, p. 62) and the dashed line plots the path of this deviation if only the permanent income shock occurred.

The lines for nondurable consumption match closely, highlighting the importance of the permanent income shock in explaining consumption. The model attributes much of the weakness in consumption during the 1990–91 recession to a decline in permanent income. It also identifies permanent factors as important contributors to spending during the boom in the second half of the 1990s. In contrast, the model estimates that most of the modest decline in consumption growth during the 2001 recession can be explained by transitory factors. Put differently, the model interprets the 1990–91 recession as being a more serious decline of the economy’s permanent productive capacity than the most recent economic downturn.

The role of permanent income shocks in the business cycle fluctuations of $x_t$ and $y_t$ is qualitatively similar: It explains some of the movements in spending on durables and total income during the 1990–91 recession and 1995–2001 boom, but is not as important in explaining the last recession. Quantitatively, however, we see that a much larger portion of the declines in $x_t$ and $y_t$ during the last two recessions were due to transitory components. Furthermore, much of the run-up in income during the second half of the 1990s was transitory. Recall that business investment boomed during the period, and the model interprets much of that increase as reflecting transitory factors.

---

**TABLE 2**

<table>
<thead>
<tr>
<th>The model's variance decompositions</th>
<th>Variance of</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$c_t$</td>
</tr>
<tr>
<td>Shock to</td>
<td>$\pi_t$</td>
</tr>
<tr>
<td>Permanent income</td>
<td>$r_t$</td>
</tr>
<tr>
<td>Others</td>
<td>$y_t$</td>
</tr>
<tr>
<td></td>
<td>$x_t$</td>
</tr>
<tr>
<td>One quarter ahead</td>
<td></td>
</tr>
<tr>
<td>Permanent income</td>
<td>62.6</td>
</tr>
<tr>
<td>Others</td>
<td>37.4</td>
</tr>
<tr>
<td>Four quarters ahead</td>
<td></td>
</tr>
<tr>
<td>Permanent income</td>
<td>73.1</td>
</tr>
<tr>
<td>Others</td>
<td>26.9</td>
</tr>
<tr>
<td>Eight quarters ahead</td>
<td></td>
</tr>
<tr>
<td>Permanent income</td>
<td>77.8</td>
</tr>
<tr>
<td>Others</td>
<td>22.2</td>
</tr>
<tr>
<td>20 quarters ahead</td>
<td></td>
</tr>
<tr>
<td>Permanent income</td>
<td>88.1</td>
</tr>
<tr>
<td>Others</td>
<td>11.9</td>
</tr>
</tbody>
</table>

13

| Does the permanent income shock say about history? | Consider figure 7. The solid line in each panel plots the deviations of the variable from the forecast based on the data in hand in 1983:Q4 (the forecasts shown in figure 5, p. 62) and the dashed line plots the path of this deviation if only the permanent income shock occurred.

14

The role of permanent income shocks in the business cycle fluctuations of $x_t$ and $y_t$ is qualitatively similar: It explains some of the movements in spending on durables and total income during the 1990–91 recession and 1995–2001 boom, but is not as important in explaining the last recession. Quantitatively, however, we see that a much larger portion of the declines in $x_t$ and $y_t$ during the last two recessions were due to transitory components. Furthermore, much of the run-up in income during the second half of the 1990s was transitory. Recall that business investment boomed during the period, and the model interprets much of that increase as reflecting transitory factors.
Forecasts

With the model’s estimates in hand, constructing forecasts of its variables is relatively straightforward. Box 2 contains a detailed description of the process. Table 3 reports the four-quarter growth rates for 2003 and 2004 and their model-based forecasts for 2005 and 2006 for nondurable consumption expenditures, privately demanded GDP, durable consumption expenditures, and the price level. These forecasts use the information available as of August 1, 2005. Recall that the model’s definitions of nondurable consumption, inflation, private GDP, and durable goods purchases are nonstandard, so the data and forecasts of table 3 are not directly comparable to the identically

Notes: The shaded areas are recessions as identified by the National Bureau of Economic Research. See the text for further details.
labeled values reported in the NIPA by the U.S. Bureau of Economic Analysis.

The model’s long-run consumption growth rate is 3.29 percent. Growth of nondurable consumption expenditures fell slightly short of this in 2003 and exceeded it by 0.40 percentage point in 2004. The forecasts of nondurable consumption growth for both 2005 and 2006 differ little from the long-run growth rate. By construction, the model’s long-run growth rate of output equals that for consumption. In fact, output growth exceeded this level by 0.50 percentage point in 2003 and fell 0.30 short of it in 2004. The model-based forecast of private GDP growth for 2005 is 0.20 percentage point below the long-run growth rate, while that for 2006 falls short by a more substantial 0.50 percentage point. The growth rate of durable goods purchases is projected to slow markedly. Given that our measure of durable goods purchases includes residential investment, this is not surprising because the growth rate of \( X_d \) substantially exceeded the model’s long-run growth rate in both 2003 and 2004. Table 3’s last row reports realized annual inflation rates and their forecasts. The model’s long-run inflation rate is 3 percent. Inflation fell below this in 2003 by 0.30 percentage point, and exceeded it by about 0.70 percentage point in 2004. The model forecasts that inflation will exceed its average in both 2005 and 2006, but by no more than 0.30 percentage point.

### The output gap

Traditional approaches to monetary policy depend on comparing the pace of economic activity to its potential. In practice, the definition of potential output growth is problematic and its measurement correspondingly difficult. Indeed, these difficulties have led Hall (2005) to advocate dispensing with the concept of potential output altogether. The structural model we present suggests one approach to this problem. We define “potential” output as the value of private GDP if the only variables affecting it were the permanent income shocks and the structural changes taken out by the 40-quarter moving averages of the two nominal ratios. The difference between this and the actual level of private GDP is our measure of the output gap. A full exploration of

### BOX 2

**Generating forecasts from the model**

Generating forecasts from the model is a two-step process. We start with the past four quarters’ values of the model’s variables, \( \Delta c, \pi, r, y, \) and \( x \). This information allows us to calculate the right-hand sides of the model’s equations. The current quarter’s values of these variables appear in the equations’ left-hand sides. We use back substitution to solve these equations and calculate our forecasts of the current quarter’s values. We denote these with \( \Delta C_{t+1}, \pi_{t+1}, r_{t+1}, y_{t+1}, \) and \( x_{t+1} \) to explicitly mark their dependence on the information available at the end of quarter \( t \). We then repeat this procedure to calculate \( \Delta C_{t+2}, \pi_{t+2}, r_{t+2}, y_{t+2}, \) and \( x_{t+2} \) to replace the unknown values of the model’s variables in period \( t \). Repeating this process generates forecasts for all of the model’s variables for any desired horizon.

This first step directly generates our forecasts for nondurable consumption growth and inflation. Because the model forecasts deviations of the ratios of private GDP and durable consumption expenditures to nondurable consumption from their 40-quarter moving averages, we must adjust the forecasts generated by the first step. Consider first \( x_{t+1} \). Because the 40-quarter moving average of this nominal ratio’s past values is available, we can account for its subtraction from the original data by adding it back to \( x_{t+1} \). That is, if we denote the original unadjusted value of the nominal ratio with \( z_{t} \), then

\[
z_{t+1} = x_{t+1} = \frac{1}{40} \sum_{j=1}^{40} z_{t-j}.
\]

To adjust \( x_{t+1} \), we use this forecast and the past 39 values of the nominal ratio to forecast the 40-quarter moving average applicable in quarter \( t + 1 \) and add the result to \( x_{t+1} \). So,

\[
z_{t+1} = x_{t+1} + \frac{1}{40} \left( \sum_{j=1}^{39} z_{t-j} \right).
\]

Adjusting the forecasts of this nominal ratio at later horizons proceeds analogously. Transforming these adjusted forecasts and the forecasts of nondurable consumption growth into forecasts of durable consumption expenditures is straightforward. The procedure for adjusting \( y_{t+1} \) and forecasting private GDP is the same.
this measure’s relationship with inflation and monetary policy lies well beyond the scope of this present article. Here, we restrict ourselves to a simple discussion of its evolution over past and current business cycles.

One can roughly gauge the path of the output gap by examining panel D of figure 7 (p. 65). Our measure of the output gap equals the vertical distance between the solid and dashed lines. Apparently, the output gap equalled zero at both of the sample’s business cycle peaks as defined by the National Bureau of Economic Research. Following the 1990 peak, the gap remained negative for about three years, underscoring the impression of a lengthy recovery from that recession. Similarly, the gap has remained negative since the 2001 peak.

An advantage of using a structural model for our forecasts is that constructing forecasts of this definition of the output gap is relatively straightforward. Figure 8 plots the actual value of the gap since 1999 as well as the model-based forecasts through the end of 2006. At the beginning of 1999, transitory shocks had driven output above what it would otherwise be by 4 percentage points. The third quarter of 2000 began a precipitous drop in the output gap. It achieved its recent trough value of approximately −3.8 percent in 2003:Q1. The gap finished our sample period at −1.2 percent. The forecast of the gap shows it rising through the end of 2006, but not by much. Its forecasted value for 2006:Q4 is −0.8 percent. By construction, the model’s forecasts of the gap approach zero as the forecast horizon increases. However, this convergence takes approximately 20 years to complete. Thus, the model implies that returning the output gap to zero or above in the near term would require some type of favorable transitory shock.

<table>
<thead>
<tr>
<th>TABLE 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>The model’s forecasts of annual growth rates as of 2005:Q2</td>
</tr>
<tr>
<td>C₁</td>
</tr>
<tr>
<td>Y₁</td>
</tr>
<tr>
<td>X₁</td>
</tr>
<tr>
<td>P₁</td>
</tr>
</tbody>
</table>

Notes: C₁, Y₁, X₁, and P₁ refer to nondurable consumption, private gross domestic product, durable goods purchases, and the price level. See the text for further details.

Conclusion

This article constructs a small macroeconomic forecasting model of real economic activity and inflation. The model identifies permanent and transitory shocks to output using only a few simple assumptions. First, the theory of permanent income tells us that real consumption of nondurables and services is uniquely informative about the long-run income prospects of the economy. Second, in the U.S. economy nominal outlays on durable goods and total private income do not persistently drift away from spending on nondurable consumption. These identifying assumptions are quite simple yet informative. The shocks to permanent income identified by the model explain most of the forecast-error variance in consumption, even at the one-quarter-ahead forecast horizon. Temporary factors explain most of the near- and medium-term forecast-error variance in spending on durable goods and total income. Nonetheless, the permanent shocks account for between one-fifth and one-third of the variation in these variables at the two-year forecast horizon, and the majority of the variance beyond 20 years.

Small-scale econometric models produce easily interpretable results that are robust to a number of specification issues which can plague large-scale systems. However, small models are too simple to address many of the issues faced by researchers and policymakers.
Accordingly, our future research will explore extensions to our model that can address interesting economic questions without adding overly restrictive identifying assumptions or burdensome model complexity.

One extension is to include the number of hours worked in the economy. Policymakers often are concerned about the implications of labor market slack for inflation, and transitory movements in hours are one proxy for such slack. Real business cycle researchers argue that changes in households’ allocation of time between work and leisure explains much of the variation in economic output. Thus, movements in hours may help forecast both inflation and output.

Second, we plan to add long-term interest rates to our analysis. Many physical assets are illiquid, and households are averse to taking on risky investment ventures. As a result, investment projects may be financed with long-term borrowing, and the interest rates on these loans will not be simple averages of current and future one-period rates. Furthermore, inflation is a key determinant of the real value of money, particularly of funds that are committed for an extended period of time. Accordingly, long-term interest rates may be a useful predictor of both real investment and inflation.

Third, some technological change is embodied in new capital machinery. For example, advances in computing power are embodied in new computer chips and improvements in fuel efficiency are embodied in new jet engines. This type of technological change will be reflected in a decline in the price of capital goods relative to the price of consumption. We plan to follow Fisher (2005) and include these relative prices, thereby allowing the model to distinguish between embodied and disembodied sources of technological progress. We speculate that this distinction will result in more informative estimates of the output gap.

Finally, following Taylor (1993), a large literature relates changes in the federal funds rate to the output gap and differences between actual inflation and the Federal Reserve’s target. We plan to investigate if the Taylor rule based on the output gap generated by our model provides any different interpretation of policy than relationships based on other measures of the output gap.
1Our model is in the spirit of Cochrane (1994). He estimates a vector autoregression in the growth rates of consumption and gross national product (GNP). The model also includes the lagged log ratio of consumption to GNP as an explanatory variable. Because the growth rates are stationary, the inclusion of the lagged log ratio forces the impact of the model’s shocks on the levels of consumption and GNP to be the same in the long run. Cochrane then identifies permanent and transitory shocks based on the restriction that shocks that do not contemporaneously affect consumption must be transitory.

2The distinction between durable goods and the service flow from those goods is subtle, and requires that we make some adjustments to the data published in the National Income and Product Accounts from the U.S. Bureau of Economic Analysis. We return to this issue in the data subsection.

3This equals the average of many independent realizations of $Z_t$. Of course, there is only one actual realization of any given variable $Z_t$. Many independent realizations can only be generated hypothetically.

4Using the unified budget constraint in equation 2 to replace the sequence of one-period budget constraints in equation 1 requires ruling out the possibility that the household finances its expenditures with a Ponzi scheme.

5Any sequence of nondurable consumption growing at this rate paired with the assumed interest rate satisfies equation 3, the condition for the optimal allocation of consumption between today and tomorrow. This growth rate, $C_p$, and the condition for optimal durable goods consumption determines a path for $S_p$, which grows at the same rate. With these paths in place, the budget constraints in equation 1 determine the path for $B_p$.

6Recall that $Y_t$ grows at the rate $\mu$ and that the one-period interest rate equals $\beta^{-1} \mu$, so the present value of the household’s labor income equals $\sum_{t=0}^{\infty} \frac{1}{(1+\beta)^t} Y_t \mu^t = \sum_{t=0}^{\infty} \beta^t Y_t = Y_0 / (1-\beta)$.

7We construct our measure of real consumption of nondurables and services as the chain aggregate of NIPA consumption expenditures on nondurables, services, and the negative of expenditures on housing services.

8The realized errors from forecasting nondurable consumption must have permanent effects on its level in order for $C_t$ to provide useful information about the economic response to changes in permanent income. In statistical terms, this means that $C_t$ must have a stochastic trend. Formal statistical tests (not shown) provide no evidence against this assumption.

9Stock and Watson (2002) review this literature.

10We later discuss how to use forecasts of these deviations to recover forecasts of the nominal ratios’ levels.

11Note that this estimator is an instrumental variables technique; the conditions in equation 9 are identical to necessary conditions identifying the lagged variables in the VAR as valid instruments.

12This forecast is constructed by taking the estimated coefficients from the model, the lagged data for 1983:Q4 and earlier, and simulating the model forward with all of the error terms set equal to zero. The VAR’s growth rate and log ratio forecasts are then transformed to the log levels shown in the figure. Box 2 (p. 66) gives more details on generating forecasts from our model.

13Cochrane (1994) finds his transitory shock is more important than the permanent shock for forecasting gross national product for about a year. After that, the permanent shock is more important.

14See Kuttner (1994) for one solution to this problem based on an unobserved components model.
REFERENCES


# Index for 2005

<table>
<thead>
<tr>
<th>Title &amp; author</th>
<th>Issue</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>BANKING, CREDIT, AND FINANCE</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Switching primary federal regulators: Is it beneficial for U.S. banks?</td>
<td>Third Quarter</td>
<td>16–33</td>
</tr>
<tr>
<td>Richard J. Rosen</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Financial constraints and entrepreneurship: Evidence from the Thai financial crisis</td>
<td>Third Quarter</td>
<td>34–48</td>
</tr>
<tr>
<td>Anna L. Paulson and Robert M. Townsend</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shopping without cash: The emergence of the e-purse</td>
<td>Fourth Quarter</td>
<td>34–51</td>
</tr>
<tr>
<td>Carol L. Clark</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>ECONOMIC CONDITIONS</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outsourcing, firm size, and product complexity: Evidence from credit unions</td>
<td>First Quarter</td>
<td>2–11</td>
</tr>
<tr>
<td>Yukako Ono and Victor Stango</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Is there evidence of the new economy in U.S. GDP data?</td>
<td>First Quarter</td>
<td>12–29</td>
</tr>
<tr>
<td>Michael A. Kouparitsas</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bringing together policymakers, researchers, and practitioners to discuss job loss</td>
<td>Second Quarter</td>
<td>2–12</td>
</tr>
<tr>
<td>Kristin F. Butcher and Kevin F. Hallock</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Henry S. Farber</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Job loss: Bridging the research and policy discussion</td>
<td>Second Quarter</td>
<td>29–37</td>
</tr>
<tr>
<td>Lisa M. Lynch</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Globalization and job loss, from manufacturing to services</td>
<td>Second Quarter</td>
<td>38–46</td>
</tr>
<tr>
<td>Lori G. Kletzer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Is retraining displaced workers a good investment?</td>
<td>Second Quarter</td>
<td>47–66</td>
</tr>
<tr>
<td>Louis S. Jacobson, Robert LaLonde, and Daniel G. Sullivan</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Understanding and addressing the challenges of job loss for low-wage workers</td>
<td>Second Quarter</td>
<td>67–74</td>
</tr>
<tr>
<td>Steven Redfield</td>
<td></td>
<td></td>
</tr>
<tr>
<td>After the doors close: Assisting laid-off workers to find jobs</td>
<td>Second Quarter</td>
<td>75–85</td>
</tr>
<tr>
<td>Randall W. Eberts</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Return on investment of high-quality outplacement programs</td>
<td>Second Quarter</td>
<td>86–93</td>
</tr>
<tr>
<td>John A. Challenger</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A lawyer’s perspective on planning a reduction in force</td>
<td>Second Quarter</td>
<td>94–107</td>
</tr>
<tr>
<td>Kenneth D. Schwartz</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Public policy and downsizing decisions</td>
<td>Second Quarter</td>
<td>108–111</td>
</tr>
<tr>
<td>Peter Cappelli</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The U.S. trade deficit: Made in China?</td>
<td>Fourth Quarter</td>
<td>2–18</td>
</tr>
<tr>
<td>Chad P. Bown, Meredith A. Crowley, Rachel McCulloch, and Daisuke J. Nakajima</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(continued)
Title & author | Issue | Pages
--- | --- | ---
**MONEY AND MONETARY POLICY**
The cost of business cycles and the benefits of stabilization  
Gadi Barlevy  
First Quarter  | 32–49
--- | --- | ---
A stable money demand: Looking for the right monetary aggregate  
Pedro Teles and Ruilin Zhou  
First Quarter  | 50–63
--- | --- | ---
Seasonal monetary policy  
Marcelo Veracierto  
Third Quarter  | 49–68
--- | --- | ---
Consumption-based macroeconomic forecasting  
Jeffrey R. Campbell and Spencer D. Krane  
Fourth Quarter  | 52–70
--- | --- | ---
**REGIONAL ISSUES**
Determinants of supplier plant location: Evidence from the auto industry  
Thomas H. Klier  
Third Quarter  | 2–15
--- | --- | ---
Energy markets and the Midwest economy  
Richard H. Mattoon  
Fourth Quarter  | 19–31
--- | --- | ---

To order copies of any of these issues, or to receive a list of other publications, please telephone (312)322-5111 or write to: Federal Reserve Bank of Chicago, Public Information Center, P.O. Box 834, Chicago, IL 60690-0834. The articles are also available to download in PDF format from the Bank’s website at www.chicagofed.org/economic_research_and_data/economic_perspectives.cfm.