Federal Reserve Bank
of New York

November 2002
Volume 8 Number 2

Economic Policy Review

Special Issue:
The Economic Effects of September 11

The *Economic Policy Review* is published by the Research and Market Analysis Group of the Federal Reserve Bank of New York. Articles undergo a comprehensive refereeing process prior to their acceptance in the *Review*. The views expressed in the articles are those of the individual authors and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System.

www.newyorkfed.org/rmaghome

Federal Reserve Bank of New York
Economic Policy Review

November 2002
Volume 8 Number 2

Special Issue:
The Economic Effects of September 11

Contents

1 Introduction
The *Economic Policy Review* Editorial Board

Articles:

5 Measuring the Effects of the September 11 Attack on New York City
Jason Bram, James Orr, and Carol Rapaport

The attack on the World Trade Center had an enormous financial, as well as emotional, impact on New York City. This article measures the short-term economic effects on the city's labor force and capital stock through June 2002, the end of the recovery process at the World Trade Center site. Using a lifetime-earnings loss concept, the authors estimate that the nearly 3,000 workers killed in the attack lost $7.8 billion in prospective income. Moreover, the employment impact in the key affected sectors—such as finance, air transportation, hotels, and restaurants—translated into an estimated earnings shortfall of $3.6 billion to $6.4 billion, while the cost of repairing and replacing the damaged physical capital stock and infrastructure totaled an estimated $21.6 billion. Accordingly, the authors determine that the total attack-related cost to New York City through June 2002 was between $33 billion and $36 billion. The article also examines the attack's effects on the city's most economically vulnerable residents and analyzes survey findings on the incidence of post-traumatic stress disorder and alcohol and drug use after September 11.

21 What Will Homeland Security Cost?
Bart Hobijn

The increased spending on security by the public and private sectors in response to September 11 could have important effects on the U.S. economy. Sizable government expenditures, for example, could trigger a rise in the cost of capital and wages and a reduction in investment and employment in the private sector, while large-scale spending by businesses could hamper firm productivity. This article attempts to quantify the likely effects of homeland security expenditures on the economy. It suggests that the total amount of public- and private-sector spending will be relatively small: the annual direct costs of the homeland security efforts are estimated to be $72 billion, or 0.66 percent of GDP in 2003. In the private sector, homeland security expenses are estimated to lower labor productivity levels by at most 1.12 percent. Therefore, the reallocation of resources associated with homeland security is unlikely to have any large and long-lasting effects on the U.S. economy.

35 When the Back Office Moved to the Front Burner: Settlement Fails in the Treasury Market after 9/11
Michael J. Fleming and Kenneth D. Garbade

Settlement fails, which occur when securities are not delivered and paid for on the date scheduled by the buyer and seller, can expose market participants to the risk of loss due to counterparty insolvency. This article examines the institutional and economic setting of the fails problem that affected the Treasury market following September 11 and describes how the Federal Reserve and the U.S. Treasury responded. The authors explain that fails rose initially because of the physical destruction of trade records and communication facilities. Fails remained high because a relatively low federal funds rate and investor reluctance to lend securities kept the cost of borrowing securities to avert or remedy a fail comparable to the cost of continuing to fail. The fails problem was ultimately resolved when the Treasury increased the outstanding supply of the on-the-run ten-year note through an unprecedented "snap" reopening. The article also suggests other ways to alleviate chronic fails, such as the introduction of a securities lending facility run by the Treasury and the institution of a penalty fee for fails.

59 Liquidity Effects of the Events of September 11, 2001

James J. McAndrews and Simon M. Potter

Banks rely heavily on incoming payments from other banks to fund their own payments. The terrorist attacks of September 11, 2001, destroyed facilities in Lower Manhattan, leaving some banks unable to send payments through the Federal Reserve's Fedwire payments system. As a result, many banks received fewer payments than expected, causing unexpected shortfalls in banks' liquidity. These disruptions also made it harder for banks to redistribute balances across the banking system in a timely manner. In this article, the authors measure the payments responses of banks to the receipt of payments from other banks, both under normal circumstances and during the days following the attacks. Their analysis suggests that the significant injections of liquidity by the Federal Reserve, first through the discount window and later through open market operations, were important in allowing banks to reestablish their normal patterns of payments coordination.

81 Has September 11 Affected New York City's Growth Potential?

Jason Bram, Andrew Haughwout, and James Orr

In addition to exacting a tremendous human toll, the September 11 attack on the World Trade Center caused billions of dollars in property damage and a temporary contraction in New York City's economy. This article explores the effect of these events on the longer run economic prospects for the city. For many years, growth in New York has taken the form of rising property prices, reflecting a steady transition from low- to high-paying jobs. During the 1990s, the city's expansion was built on several factors, including improving fiscal conditions, better public services, and shifting industrial and population structures that favored job and income growth. The study suggests that the effects of September 11 will not eliminate these advantages in the medium term; in fact, preliminary indications are that the city remains an attractive location for businesses as well as households. Nevertheless, New York City will face many challenges as it attempts to return to its pre-attack growth path.

97 Terrorism and the Resilience of Cities

James Harrigan and Philippe Martin

The September 11 attacks in New York and Washington have forced Americans to confront the fact that to live or work in a large city is to be at greater risk of large-scale terrorism. What do these risks, and the public perception of them, imply for cities in general and the future of New York City in particular? In this article, the authors begin their exploration of this issue by examining why cities exist in the first place. To conduct their analysis, they simulate two key theoretical models of economic geography, using data that approximate the characteristics of a major U.S. city as well as estimates of the costs of the September 11 attacks. The authors conclude that the very forces that lead to city formation also lead cities to be highly resilient in the face of catastrophes such as terrorist attacks. They argue that New York City in particular is likely to continue to thrive despite any ongoing terrorist threat.

Introduction

This special volume of the *Economic Policy Review*, issued soon after the one-year anniversary of the attacks of September 11, 2001, explores some of the key economic consequences of the attacks. The six articles that make up the volume address several important questions: How great were the losses in New York City on September 11 and in the difficult months thereafter? How much will the nation spend to prevent future attacks? Did the destruction of information and infrastructure impair the functioning of the payments and securities settlement systems, and what steps minimized further damage? Will these events hurt New York's future vitality and cause businesses and workers to retreat from the city?

The authors consider these questions from their vantage point in the Research and Market Analysis Group of the Federal Reserve Bank of New York. Economists in the Group conduct research and policy analysis in support of the New York Fed's responsibilities, which include carrying out the Federal Reserve System's open market and foreign exchange operations, managing payments-clearing activities between banks, contributing to the formulation of monetary policy, and tracking economic conditions in New York City and the Second Federal Reserve District. The economists' knowledge and experience in these areas inform their analysis of the September 11 events, as does their strong commitment to understand and illuminate events that took place in the Bank's own Lower Manhattan neighborhood.

The articles in the volume range in subject matter—from the concrete effects of the attacks on the financial system infrastructure to more abstract issues such as the viability of cities in the face of terrorism. But while the articles touch on a wide variety of topics, all employ the language and the viewpoint of economics. Thus, the authors follow the conventions of their discipline when they use forgone earnings to measure the "cost" of lives lost in the attack on the World Trade Center. Certainly, these losses should not be omitted from accountings, and forgone earnings is the best measure available for this purpose. However, the authors and editors recognize very clearly that the true cost of September 11—the grief, terror, and general loss of well-being—is beyond any calculus.

Overview

The six articles in the volume fall into three broad groups: 1) detailed accountings of economic costs—those incurred as a direct consequence of the September 11 attacks and those arising from efforts to prevent future attacks, 2) studies of the attacks' disruptive effects on the payments and securities settlement systems, and 3) analyses of New York City's prospects after September 11.

Cost Accountings

Jason Bram, James Orr, and Carol Rapaport begin the volume by analyzing the costs of the attack on New York City. Their account is indeed sobering. New York lost an estimated $7.8 billion in prospective earnings from those who died in the attack and $3.6 billion to $6.4 billion in job earnings from workers displaced afterward. These losses are distinct from the concurrent income losses attributable to the national recession and are spread across both high- and low-wage industries. Property losses—including the destruction of real estate, subway lines, and communications equipment—and the costs of cleanup and site restoration are expected to reach

$21.6 billion. Adding up the earnings losses, destruction of property, and cleanup costs, the authors calculate the direct damage from the attack on the World Trade Center to be between $33 billion and $36 billion. The additional productivity losses from psychological stress are harder to estimate, but quite real—as reflected in several surveys of city residents' experiences. Although insurance payments and federal aid have helped markedly to mitigate the attack's financial impact on New York City residents and businesses, these financial transfers do not lessen the more personal and psychological effects of the destruction.

The authors are careful to note that their cost estimates reflect the information available as of June 2002, the end of the recovery process at the World Trade Center site; the actual magnitude of the costs will become much clearer over time. There are, however, some signs that the brunt of these costs has already been borne: the Trade Center site has been cleaned up ahead of schedule, employment levels appear to have bottomed out, and housing market indicators remain surprisingly buoyant.

Focusing on the national effects of the attack, *Bart Hobijn* assesses how much the country might spend to prevent more such incidents. The federal government plans to spend $38 billion in 2003 on border security, protection against biological threats, emergency preparedness, and various other homeland security measures. State and local governments are expected to spend about $1.3 billion on security next year. Although a public expenditure of $39 billion is large in absolute terms, it is small relative to the $10 trillion expected output of the U.S. economy next year. Private-sector spending on counterterrorism is harder to gauge. Adopting the assumption that firms will double the amount they have spent in recent years on "protective services" (security guards, surveillance equipment, and so forth), Hobijn estimates that private security spending next year will reach about $33 billion. Thus, the author's total figure for the *direct costs* of homeland security spending for fiscal year 2003—private and public—is about $72 billion.

The effort to protect the nation from further attacks also entails indirect costs—including the inconvenience and delays experienced by the public as heightened security measures are put into place. Hobijn focuses on the costs associated with increased waiting times at airports. Drawing on a variety of sources, he estimates that about $12 billion will be lost this year in production and leisure time because passengers are waiting longer to pass through security checkpoints at airports.

Signficantly, Hobijn does not try to evaluate the efficacy of the $72 billion expenditure in securing the country against further attacks. A complete accounting of the economic costs of the attacks—one combining the $33 billion to $36 billion estimate advanced by Bram, Orr, and Rapaport for New York City and the costs associated with the Pentagon attack and the Pennsylvania crash—can only suggest the size of the loss that the country is seeking to prevent. A proper evaluation of the $72 billion expenditure would require an assessment of its impact on the likelihood of future attacks and on intangibles such as peace of mind—a very different, and difficult, question.

The Impact on the Payments and Securities Settlement Systems

The World Trade Center was not only an important symbol of business and finance, it was also a key location for those activities. The attack claimed lives, destroyed physical capital, and disrupted the information flows that facilitate transactions in financial markets. How well did the U.S. financial system withstand the blow? In particular, how did the various payments mechanisms—that is, the facilities used to transfer large amounts of money and financial instruments among institutions—perform on September 11 and the days that followed? And how did important market participants and policy institutions, including the Federal Reserve and the Treasury Department, respond?

Two articles in the volume examine the problems that arose in the payments and securities settlement systems in the wake of the September 11 attacks. The first studies the prolonged failure of brokers, dealers, and investors to deliver on Treasury security trades, and the second examines the disruption to Fedwire, the electronic network that processes large payments among financial institutions. The articles describe the technical exigencies and market conditions that led the Federal Reserve to inject vast amounts of liquidity into the banking system and to relax normal limits on securities lending.

Michael Fleming and Kenneth Garbade set the stage for their study of Treasury market functioning with a review of the market's normal settlement procedures. They then describe how the events of September 11 led to a huge and prolonged increase in settlement "fails" as sellers of Treasury securities did not meet their commitments to deliver securities on the dates scheduled. During the week ending September 19, daily average fails increased to $190 billion, a sharp rise from the $7.3 billion daily average observed during the first eight months of 2001. The increased fails can be traced to the technical problems created by the massive physical destruction in Lower Manhattan: several brokers and dealers could not operate, a clearing bank for dealers had to close its downtown offices, telephone lines were severed, and trading records were

lost. The Federal Reserve responded by relaxing its limits on securities lending, thereby making Treasury collateral more readily available. This action, together with restored communications, made the situation less acute, but fails persisted at higher than usual rates into October.

How did severe, but relatively short-lived, technical problems cause an extended disruption? Fleming and Garbade explain that "specials" rates—the rates that investors earn when they lend cash to borrow the particular securities they need to settle earlier trades—were near zero for several actively traded securities. The low specials rates gave investors little incentive to borrow securities to avert or remedy fails. The authors attribute the low specials rates in the weeks after the attack to a low federal funds rate and to concern among owners of securities that lent securities might not be promptly returned. Ultimately, worries that chronic fails were undermining market functioning led the Treasury to reopen its ten-year note. The additional supply raised the note's specials rate and increased investors' incentives to settle their trades. Fleming and Garbade conclude with a discussion of longer run reforms to prevent chronic fails, including the creation of a Treasury facility that could lend specific securities on a temporary basis and the imposition of a penalty fee for fails.

James McAndrews and Simon Potter investigate the problems affecting Fedwire, a backbone of the U.S. payments systems. Payments fell off sharply after the attacks: instead of the usual $10 billion per minute that flows over the wire near the end of most days, the end-of-day flow on September 11 was less than $2 billion per minute. The immediate problem was logistical: some banks were unable to send payments because of the destruction caused by the attacks. Because banks routinely use anticipated receipts to fund their own payments, many banks expecting payments found themselves unexpectedly short of liquidity. As a consequence, these banks were less likely to send payments to other banks in the normal pattern.

As the coordination of payments broke down, upsetting the distribution of balances across the banking system, banks sharply increased their precautionary demand for liquidity. McAndrews and Potter track this phenomenon by estimating banks' payments responses to the receipt of payments from other banks. The authors' analysis shows a distinct shift after September 11, with the typical bank requiring more of a liquidity cushion in advance of sending out payments than it had in the preceding period.

The Federal Reserve responded to the payments disruption by supplying extraordinary amounts of liquidity to the banking system: it loaned billions of dollars through the discount window (more than two hundred times the daily average amount of lending in the prior month), temporarily suspended penalties for bank overdrafts, and bought securities on the open market. In addition, Federal Reserve staff contacted the banks to assure officials of the availability of discount loans and to encourage them to make payments as usual. These actions by the central bank, McAndrews and Potter argue, helped reestablish payments coordination.

The authors identify the discount window as a particularly valuable tool in restoring coordination after September 11. The effectiveness of discount window loans calls into question the view espoused by some that open market operations alone can meet the liquidity needs of banks in extreme circumstances. The authors also discuss longer run payments system reforms—including infrastructure changes and changes in the protocols for submitting and settling payments—that might preserve coordination in the event of future disruptions.

Prospects for New York City: Can the Center Hold?

The September 11 attack was a severe blow to New York City. Could this event jeopardize the standing of the city as a leading financial capital? Will the trends toward higher income and improved quality of life in the city be halted or reversed by the attacks? Interestingly, the two articles that assess the city's longer term prospects reach similarly optimistic conclusions, despite very different approaches. One examines the economic trends that have helped New York City to prosper in recent decades and considers whether the attack will force the city off its course. The authors' findings indicate that New York's mix of industries should continue to serve the city well over the near horizon. The second article looks at current theories about why cities come into being, in order to assess the likelihood that terrorism could threaten the existence of New York. Here, the analysis suggests that forces strong enough to create cities are very difficult to overcome by terrorist actions.

Jason Bram, Andrew Haughwout, and James Orr use data on rents, wages, and labor shares to argue that the economic prospects for New York City at the time of the attack, even with the incipient recession, were favorable. Over a period of twenty-five years, New York City has enjoyed stable employment, rising real earnings, and appreciating land prices. Earnings in the city have advanced at a rate well above rates in the rest of the country, owing to accelerating productivity in the city's existing jobs and expanding employment in the high-paying services sectors—most notably finance. The high rents in New York, though lamented by residents, offer a clear indication that people are willing to pay a premium to live in

the city. Together, these patterns are consistent with a model of New York as an attractive, mature city in an open, competitive environment.

Looking forward, the authors tie the city's future to the supply of amenities such as safety, the arts, and municipal services, and to the demand for the goods and services produced in New York City. The authors find that trends in these areas have been encouraging in recent decades. With respect to amenities, crime rates have fallen rapidly, and the city government has strengthened New York's fiscal position, lowered property taxes, and improved public transit. Whether the fiscal strains from the attack will reverse these gains, the authors argue, will depend on how the city manages its finances and how it rebuilds the destroyed infrastructure.

As to the demand for its goods and services, New York City has a high concentration of some of the nation's fastest-growing industries, suggesting that the city's specialties tend to be in high demand. Moreover, national employment projections indicate that over the next decade, most of the city's key industries should fare better than average in generating jobs.

Seen in this context, the terrorist attack would be most damaging if it upset either the amenities available to city residents and businesses, or the demand for New York's products and services. Preventing such outcomes, the authors suggest, is the challenge now facing the city. New York's policymakers will need to close a substantial city budget gap without letting crime rise or municipal services deteriorate significantly, and without pricing New York's products and services out of the market. Judicious policy choices, along with the support of federal aid, will be key to New York's economic growth.

James Harrigan and Philippe Martin provide a more abstract analysis of New York City's prospects. They assess the viability of cities, and New York in particular, in the face of catastrophes such as terrorist attacks by considering why cities exist in the first place. The authors draw on two models, or theoretical explanations, for the existence of cities. The first centers on the idea that cities "pool" labor (that is, offer workers and firms an easy way to find each other); the second is based on the notion that cities lower transport costs (for goods shipped between producers and consumers). These rationales for the existence of cities are called agglomeration forces. In both models, the stable outcome—a city in equilibrium—is very stable indeed, because the agglomeration forces that create the city also tend to preserve it.

Using these models, Harrigan and Martin ask whether terrorism could overcome a city's agglomeration forces, causing firms and workers to scatter and the city to decline. They describe terrorism and the threat of terrorism as a special type of "tax" on a city's firms or residents, reflecting the costs of such hardships as higher insurance rates and security-related delays. This new tax detracts from firms' profits or workers' income without funding improvements in infrastructure or services, as a normal tax might.

The authors' simulations of the two models, conducted with data that approximate the characteristics of a large U.S. city, suggest that the vitality of cities could withstand terrorism "tax" rates well in excess of those that are likely to occur. Given the magnitude of the economic benefits that major cities generate, New York and its counterparts elsewhere should be remarkably robust in the event of subsequent terrorist attacks.

Concluding Remarks

The authors and editors hope that this volume will contribute to a fuller understanding of the September 11 events and their aftermath. In calculating the costs of the New York attack and those of protecting the nation from further assaults, the volume provides a measure of the injury sustained by the United States. But while these cost estimates may in some sense speak to the country's vulnerability, many of the volume's findings underscore the strong performance of our markets and institutions—both national and local—in a time of crisis. Economic data presented here show that New York City's prospects for growth remain favorable, while the economic theory outlined in the collection's final article upholds the resilience of all cities in the face of great shocks. The articles detailing events in the Treasury market and the nation's payments system tell a similar story of recovery, affirming the flexibility of U.S. financial and regulatory institutions and the resourcefulness of the individuals within them.

Erica L. Groshen, Linda S. Goldberg, James A. Kahn, Hamid Mehran, Donald P. Morgan, and Carol L. Osler

Bart Hobijn

What Will Homeland Security Cost?

- As the government and the private sector increase their spending on security measures, concerns have arisen over the magnitude and economic effects of these expenditures.
- A review of the evidence, however, suggests that public- and private-sector outlays for increased security will be relatively small—roughly $72 billion per year, or 0.66 percent of the nation's GDP in 2003.
- Fears that private-sector productivity will decline significantly as firms shift resources to protection appear ill-founded. Firms' security initiatives will lower labor productivity levels by no more than an estimated 1.12 percent.
- Indirect costs of homeland security—notably, the delays related to heightened airport security and the diversion of research and development funds from productivity-enhancing technologies—are also likely to prove modest.

Increased spending on security measures—high on the agenda of the government and much of the private sector—will undoubtedly have an effect on the economy. If firms devote sizable amounts of time and money to the protection of their businesses, they will reduce their overall productivity. Furthermore, if the government's spending on homeland security is significant, it could lead to a rise in the cost of capital and wages and reduce investment and employment in the private sector. Finally, the homeland security efforts could have many indirect economic effects, such as the costs of increased airport waiting times and long-run productivity effects resulting from a reallocation of research and development (R&D) spending.

In this article, we attempt to quantify the economic effects of the homeland security efforts of the public and private sectors, focusing specifically on the costs of these efforts. In practice, it is difficult to classify which expenditures are related to homeland security. For this reason, we use the broadest possible definition: all expenditures *possibly* aimed at either preventing damage due to terrorist attacks or at preparedness for the response to potential attacks. This broad definition suggests that the figures presented here should be interpreted only as estimates of the maximum effect of homeland security on the economy.

From this perspective, our study can be interpreted as the cost side of a cost-benefit analysis. To estimate the costs of homeland security, we focus on three main questions. The first involves the

Bart Hobijn is an economist at the Federal Reserve Bank of New York.

The author thanks two anonymous referees. The views expressed are those of the author and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System.

likely magnitude of government expenditures on homeland security. To answer this question, we review the historical as well as the proposed expenditures on homeland security by the federal, state, and local governments and compare them with historical spending on other programs and items.

The second question concerns whether firms will spend significant amounts of time and money on security and protection. To offer insight into this issue, we estimate the share of inputs that firms devoted to protective services before 2001 and consider what effect doubling this share would have on firm productivity levels. This technique is similar to the one used by the Council of Economic Advisers (CEA, 2002) to assess the productivity effect of homeland security.

Despite our broad definition of homeland security expenditures, our results suggest that the amounts of public- and private-sector spending are likely to be relatively small.

The final question involves the size of the many indirect effects of homeland security. Of the many possible effects, we examine two in particular: the costs associated with increased waiting times at airports and the possible effect on long-run productivity growth attributable to a shift in R&D expenditures. The answer to this question turns out to be the most speculative of the three.

Despite our broad definition of homeland security expenditures, our results suggest that the amounts of public- and private-sector spending are likely to be relatively small: the total annual direct costs of the homeland security efforts are estimated to be $72 billion, or 0.66 percent of GDP in 2003. Moreover, the homeland security efforts in the private sector are estimated to lower labor productivity levels by at most 1.12 percent. Consequently, the reallocation of resources due to homeland security is unlikely to have any large and long-lasting effect on the economy. Furthermore, spending on homeland security should not evaporate the "peace dividend" of the 1990s (that is, push defense expenditures back to their Cold War levels and force the federal government to run large budget deficits). Even when we include homeland security spending, the proposed defense budget will still make up a smaller fraction of GDP than it did in any year from 1947 to 1994.

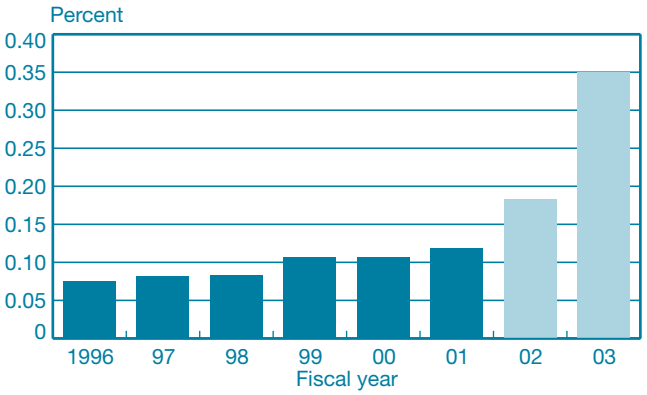
Spending on Homeland Security

Public Sector

The term homeland security was only introduced formally after the September 11 terrorist attacks. However, the federal government had previously been funding several anti-terrorism programs, many of which span several agencies. In 1995, the National Security Council was assigned to coordinate these programs and the Office of Management and Budget (OMB) was appointed to supervise their budgetary aspects. In practice, however, most agencies did not specifically account for expenditures as being "terrorism-related" until 1998.[1] The 1998 National Defense Authorization Act requires each administration to provide an annual report on the funding of programs to combat terrorism. Since 1998, OMB has provided Congress with an annual overview of terrorism-related expenditures, which include funds to combat terrorism, to prepare for a response to weapons of mass destruction, and to protect critical infrastructure.[2]

The federal government's recent expenditures on anti-terrorism/homeland security as a percentage of GDP are presented in Chart 1. In the six years before September 2001,

Chart 1
Federal Anti-Terrorism/Homeland Security Expenditures as a Percentage of GDP



Sources: U.S. Department of Commerce, Bureau of Economic Analysis, National Income and Product Accounts data; 1996 and 1997 data: General Accounting Office (1997); 1998-2001 data: Office of Management and Budget (2001); projections for 2002 and 2003: House Budget Committee and White House press releases; GDP projections for 2002 and 2003: Congressional Budget Office testimony (2002).

Note: Fiscal years run from fourth quarter to third quarter.

the government had spent about 0.1 percent of GDP on its anti-terrorism program. Immediately before September 2001, the approved spending to combat terrorism was $12 billion for fiscal year 2002, which again would have been about 0.1 percent of GDP. In response to the attacks, however, Congress supplemented the budget, allowing the Bush Administration to spend approximately $20 billion on homeland security.

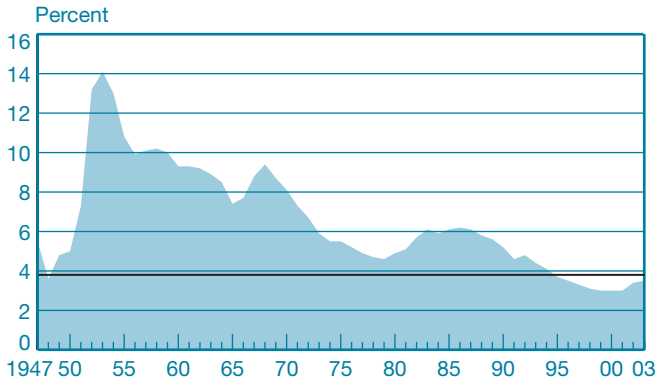
In the February 4, 2002, budget proposal, the President's current budget sets aside $38 billion for homeland security expenses for the 2003 fiscal year. Projections by the Congressional Budget Office (2002) indicate that this amount would represent 0.35 percent of GDP. Consequently, in response to the terrorist attacks, the federal government has tripled the amount of its homeland security expenditures as a share of GDP.

Accordingly, it is reasonable to question the size of homeland security expenditures relative to other budget items. Among these many items, defense expenditures offer the best basis for comparison. A major concern is that the additional expenditures on homeland security will erase the peace dividend of the 1990s (see, for example, *BusinessWeek* [2001] and Baily [2001]). Of the $38 billion earmarked by the Bush Administration for homeland security in fiscal year 2003, $7.8 billion is part of the defense budget. The proposed total defense budget for 2003 is $379 billion and will represent about 3.5 percent of GDP. Hence, if one were to add to the defense budget the additional $30 billion in nondefense homeland security spending, homeland security expenditures and national defense outlays would account for about 3.8 percent of GDP. This share would be about the same as the share of national defense outlays in 1995 and would still be lower than it was in any year between 1947 and 1994 (Chart 2). Therefore, the concern that homeland security will eliminate the peace dividend of the 1990s seems unfounded, at least for the currently proposed spending levels.

The Bush Administration plans to devote its homeland security budget to five main objectives: support of first responders in preparation for future terrorist threats, improvement of the U.S. response to biological terrorism, improvement of border controls, tightening of aviation security, and enhancement of information-sharing on potential terrorists (Chart 3).[3] Bear in mind that many of these objectives involve significant subsidies to state and local governments to support their efforts to prepare for and prevent possible attacks.

The overall expenditures that state and local governments will incur due to increased security measures in response to the September 11 attacks are less transparent than those of the federal government. This lack of transparency could be a source of delays in local security efforts (*New York Times* 2001). Although we do not know how much state and local governments will allocate to homeland security, two surveys shed light on their possible expenditures. A survey by the National Governors Association, conducted in December 2001,

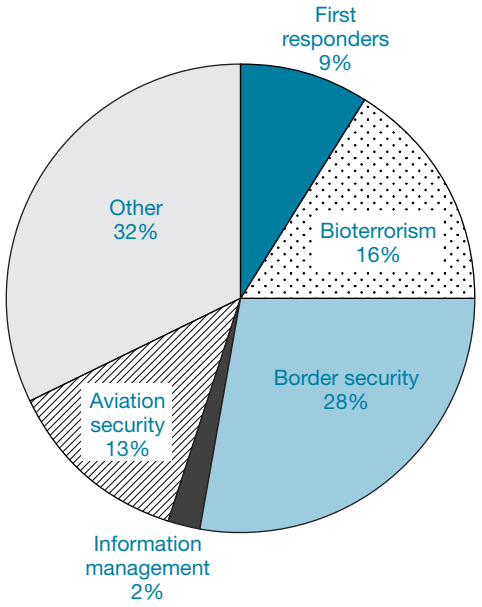
Chart 2
Share of National Defense Outlays in GDP



Sources: Congressional Budget Office (2002); author's calculations.

Note: The shaded area represents defense outlays as a percentage of GDP; the solid line represents the 3.8 percent of GDP that is the share of the combined defense and homeland security budgets in fiscal year 2003.

Chart 3
Composition of 2003 Homeland Security Budget



Source: Office of Management and Budget (2002).

suggests that state governments expect the costs of homeland security for 2002 to be as high as $4 billion. Of this amount, $3 billion would be required to improve emergency communications systems and bioterrorism preparations, while the remainder would be used for critical infrastructure protection. A survey by the U.S. Conference of Mayors, in January 2002, indicates that additional security costs for all cities with a population of 30,000 or more could be as high as $2.1 billion for 2002.[4] By comparison, the City of New York forecasts $3.8 billion in spending on its police department in 2002.[5]

The concern that homeland security will eliminate the peace dividend of the 1990s seems unfounded, at least for the currently proposed spending levels.

Thus, by combining the results of these surveys, we arrive at a preliminary estimate of about $6.1 billion to fund homeland security efforts on a state and local level. However, there is a large overlap between this amount and the federal homeland security budget. The federal budget includes $4.8 billion to support local first responders and bioterrorism preparedness. Consequently, the net expenditures of state and local governments on homeland security will likely be about $1.3 billion, which would be a very small expense when compared with the $1,276 billion in state and local government expenditures in 2001. Furthermore, the additional security expenditures of state and local governments might be lower going forward, because their initial expenditures largely represent one-time investments in necessary equipment. In fact, the U.S. Conference of Mayors suggests that 50 percent of additional security expenditures by cities are capital expenditures.

These figures suggest that homeland security expenditures are relatively small compared with those of other government programs. Thus, those who argue that fiscal discipline should have a high priority in response to September 11, as Baily (2001) does, need not be concerned about homeland security expenditures having a major effect on the budget.[6]

Private Sector

Besides an expansion of the public sector, another concern is that the private sector will devote a large part of its time and resources to protective rather than productive activities. Such a shift in resources would raise unit production costs and lower productivity. Although we cannot estimate exactly how much money and time firms will allocate to increased security, we can measure how much they have spent in the past on security initiatives. After determining that amount, we can consider the effect on firm productivity of a hypothetical doubling of these resources.

There is no separate accounting of firms' expenditures on anti-terrorism security.[7] The closest we can get to measuring these expenditures is to consider the much more general classification of "protective services." Box 1 describes the parts of labor, capital, and intermediate inputs classified as protective. Clearly, protective services encompass much more than terrorism-related security. They vary from fire protection to the protection provided by crossing guards. This classification of all protective services inputs as related to homeland security suggests that our results are best interpreted as an estimate of the maximum effect of homeland security on the private sector.

Box 1
Protective Services in the Private Sector

- **Labor: Protective Services Occupations**[a]
  Category 33 of Standard Occupational Classification; mainly firefighters, police officers, correctional officers, private detectives, and security guards.
- **Capital: Electronic Security Systems**[b]
  Electronic access control, anti-burglary, closed-circuit television, fire protection, systems integration, and home automation.
- **Services: Protective Services Industry**[c]
  Establishments engaged primarily in providing one or more of the following: 1) investigation and detective services, 2) guard and patrol services, 3) pick up and delivery of money, receipts, or other valuable items, with personnel and equipment to protect such properties while in transit.

[a]Source: U.S. Department of Labor, Bureau of Labor Statistics, Occupational Employment Survey.

[b]Sources: Security Industry Association of America, *Research Update; Security Sales Magazine.*

[c]Source: U.S. Department of Commerce, Bureau of Economic Analysis, Input-Output Tables.

Furthermore, there are areas not included in the list of protective services inputs to which firms might devote more resources in response to a heightened threat of terrorism. For instance, the capital measure includes only electronic security systems and not items such as fencing. However, Anderson's study (1999) of the costs of crime in the United States suggests that the value of these additional types of capital is small relative to that of the electronic security systems included in the analysis. In addition, the data do not account for time spent by nonprotective services employees on terrorism preparedness. Finally, the measures of security-related inputs used here do not account for the increased information security measures that firms will take to protect their computer networks and resources, nor do they include the expenditures that firms incur by investing in the establishment of operational and informational back-up sites.

Although there is no direct evidence on these unaccounted-for inputs, some preliminary findings exist. O'Hanlon et al. (2002) suggest that the total annual homeland security cost for the private sector will be about $10 billion; this amount includes the security measures accounted for here as well as the unaccounted-for items. More specifically, data presented in a report by RBC Capital Markets (2001) indicate that before September 11, prepackaged security software made up about 2 percent of firms' software expenditures, and security-related computer equipment represented about 0.3 percent of total computer equipment expenditures. Total Internet security expenditures made up a larger part of inputs because these data do not take into account expenditures on custom-made and in-house software applications, personnel assigned to Internet security-related activities, and computation time devoted to software protection rather than to other applications.

To evaluate the importance of protective services workers in private employment, we consider two measures. The first is the share of workers in protective services (Table 1). Here, we see that the 2000 share of these personnel in private employment is about 1.12 percent.[8] The approximately one million security guards account for about 80 percent of these workers in the private sector. What is relevant for measuring the effect of these workers on productivity, however, is not only their share in

Table 1

Share of Inputs Devoted Directly to Protective Services

Percent

Sector	Employment[a]	Wage Bill[a]	Replacement Value of Capital[b]	Capital Service Flows[b]	Intermediate Services
Business	1.12	0.71	0.52	0.46	—
Nonfarm business	1.13	0.71	0.54	0.44	—
Manufacturing, total	0.16	0.13	0.13	0.04	0.76
Durables	0.16	0.13	0.15	0.05	0.92
Nondurables	0.16	0.14	0.12	0.04	0.60
Government	17.26	17.53	0.38	—	—
Federal	2.94	2.98			
Police officers	0.63	0.52			
Security guards	0.13	0.08			
State[c]	17.04	16.55			
Police officers	2.15	2.49			
Security guards	0.29	0.19			
Local	25.09	28.32			
Police officers	10.31	12.00			
Security guards	0.25	0.17			

Sources: U.S. Department of Labor, Bureau of Labor Statistics, Occupational Employment Statistics (2000); U.S. Department of Commerce, Bureau of Economic Analysis, Fixed Assets Tables; *Security Sales Magazine*; Security Industry Association of America, *Research Update* (2001, second quarter).

Note: Data are constructed for 1999.

[a]Labor input data are for 2000 and for all workers in protective services occupations.

[b]Protective capital is assumed to consist of electronic access control, anti-burglary, closed-circuit television, and fire protection systems, as well as systems integration and home automation.

[c]Correctional officers represent the majority of protective services individuals working for state governments.

employment, but also their share in the total wage bill. Since protective services employees tend to earn an hourly wage that is about 63 percent of the average, their share in the total private wage bill is lower than their share in total employment. More precisely, these workers earn only 0.71 percent of the wage bill; by comparison, lawyers account for only 0.32 percent of the labor force, but earn 0.96 percent of the wage bill.

If businesses would have had to replace their capital stocks in 1999, then only about half of 1 percent of that replacement cost would have been spent on electronic security systems. Moreover, of the portion of 1999 output that can be accounted for as produced by capital inputs, only 0.46 percent was attributable to these systems. In technical terms, these figures imply that the share of electronic security systems in the total private capital service flows is 0.46 percent.

To estimate the effect of homeland security on private-sector productivity, we now consider a scenario in which this sector doubles the security-related capital and labor inputs to productivity.[9] This scenario is similar to the one employed by the Council of Economic Advisers (2002). The effect of such a scenario depends on the type of productivity considered. The Bureau of Labor Statistics measures two types of productivity: labor productivity, the amount of output produced per hour worked, and multifactor productivity (MFP), the amount of output produced per unit of input, where a unit of input is measured as a combination of labor and capital. Box 2 describes the productivity concepts applied in this study and the productivity calculations performed.

Box 2
Calculations of Productivity Effects

The Bureau of Labor Statistics measures two types of productivity. Here, we describe how the homeland security efforts can affect both measures.

Notation and Assumptions

Δy is the growth rate of output,
Δl is the growth rate of labor inputs,
Δk is the growth rate of capital inputs.

The main assumption is that firms will double their security-related capital and labor inputs while output levels remain the same. That is, $\Delta y = 0$, while Δl is the share of security-related labor inputs in overall labor inputs and Δk is the share of electronic security systems in the total capital input level (capital service flows).

Labor Productivity (ALP)

Definition: Labor productivity is the amount of value-added output produced per hour worked.

Effect: The growth rate of average labor productivity is the difference between the growth rate of output and that of the labor input, that is, $\Delta ALP = \Delta y - \Delta l$. For the scenario considered here, this implies that ΔALP is minus the share of security-related labor inputs in the overall hours.

Multifactor Productivity (MFP)

Definition: Multifactor productivity is the amount of output, measured as value added, produced per normalized unit of inputs. Units of inputs are measured as a combination of capital, that is, equipment and structures, and labor. Labor is measured in wage-adjusted hours. Wages are adjusted based on the assumption that workers who make a wage that is twice as high would be twice as productive, which should be adjusted for in accounting for the hours that they work.

Effect: The growth rate of multifactor productivity equals the growth rate of output minus a weighted average of the growth rates of the capital and labor inputs, that is, $\Delta MFP = \Delta y - w\Delta l - (1 - w)\Delta k$, where the weight w is determined by the nominal output share of labor. For the scenario considered here, this implies that $\Delta MFP = -w\Delta l - (1 - w)\Delta k$.

Multifactor Productivity (Capital, Labor, Energy, Materials, and Services) (MFP-KLEMS)

Definition: MFP-KLEMS is the amount of gross output produced per normalized unit of inputs. Units of input are measured here as a combination of capital, labor, energy, materials, and services. The Bureau of Labor Statistics measures labor here in terms of hours worked.

Effect: The same as multifactor productivity, with the additional effect that firms might also increase the business services they buy to cover some of their security activities. Therefore, in this case, homeland security measures affect productivity through the capital and labor input channels, as well as through business services inputs.

Notes: The value-added concept of output refers to total output minus the value of all the intermediate inputs a firm buys. See U.S. Department of Labor, Bureau of Labor Statistics (1997, Chapters 10, 11) for more details on productivity definitions and statistics. Stiroh (2001) presents a more formal discussion of some of the terms involved.

Table 2

Effect of Homeland Security Efforts on the Level of Productivity

Percentage Change

Scenario	Type of Productivity	Business	Nonfarm Business	Manufacturing	Durables Manufacturing	Nondurables Manufacturing
Doubling of security-related labor inputs	ALP	-1.12	-1.13	-0.16	-0.16	-0.16
	MFP	-0.49	-0.49			
	MFP-KLEMS			-0.06	-0.07	-0.05
Doubling of security-related capital inputs	ALP	0.00	0.00	0.00	0.00	0.00
	MFP	-0.14	-0.15			
	MFP-KLEMS			-0.01	-0.01	-0.01
Doubling of security-related business services inputs	ALP					
	MFP					
	MFP-KLEMS			-0.10	-0.11	-0.08
Additional airport delays	ALP		-0.09			
	MFP		-0.08			
	MFP-KLEMS					
Total effect	ALP	-1.12	-1.22	-0.10	-0.11	-0.09
	MFP	-0.63	-0.71			
	MFP-KLEMS			-0.17	-0.19	-0.14

Source: Author's calculations.

Note: ALP is labor productivity; MFP is multifactor productivity; MFP-KLEMS is multifactor productivity (capital, labor, energy, materials, and services).

Our calculations of homeland security's effect on productivity assume that businesses increase their inputs by doubling the resources devoted to their protective activities. These inputs are assumed to be unproductive, however, in the sense that they do not lead to any measurable output. Table 2 reports the estimates of doubling various inputs to labor productivity as well as to multifactor productivity. Note that all of the effects presented are effects on the productivity *level* and not on the growth rate.

The largest productivity effect from doubling all security-related labor inputs would be the effect on labor productivity in the nonfarm business sector: we estimate that productivity would be lowered by about 1.13 percent. The effect is lower for MFP because the MFP measure takes into account the fact that protective services personnel earn below-average wages, and it adjusts for the labor share. Doubling security-related labor inputs would lower multifactor productivity only by about half of 1 percent. This effect is still most likely an overestimation of homeland security's actual effect, not only because we are assuming that the private sector would double the number of protective services workers, but also because all of these workers are not protecting private-sector businesses; some are providing security services to the government and consumers. Furthermore, many of these workers guard against incidents other than potential terrorist attacks.

Next, doubling the number of installed electronic security systems would lower multifactor productivity only by an estimated 0.15 percent in the business sector, and it would have virtually no effect on manufacturing productivity levels. Finally, the effect of doubling security-related service inputs to manufacturing would lower MFP by approximately 0.1 percent.

In sum, doubling all inputs directly related to security in the business sector would most likely lower MFP by 0.8 percent or less and would lower labor productivity by at most 1.12 percent. This estimated effect is of a magnitude similar to the effect reported by the Council of Economic Advisers. The CEA estimates that the homeland security efforts will reduce output over five years by 0.6 percent relative to its level without the efforts. Note that the CEA's calculation does not only include the shift in productivity estimated here. It also includes the effect of a decrease in expected investment—and therefore a decline in the future capital stock—due to this downward shift

in productivity. It is important to realize, however, that it might take some time before the productivity effect calculated here is realized: firms need time to implement additional security measures.

Although the scenario depicted here is comparable to the one offered by the CEA, it is still reasonable to ask how likely it is that firms will actually double security-related inputs. One way to answer this question is to compare the cost of this doubling with the $10 billion in private-sector homeland security costs estimated by O'Hanlon et al. (2002). Doubling the number of protective services workers in the business sector alone would already cost more than $25 billion annually at

Our calculations of homeland security's effect on productivity assume that businesses increase their inputs by doubling the resources devoted to their protective activities.

2000 wage levels. If we add to this amount the $102 billion replacement value of the electronic security systems in place in the United States, measured in 2000 dollars, we see that the implementation costs of doubling the security-related inputs vastly exceed the estimates of O'Hanlon et al.[10] Thus, our estimates are best interpreted as an upper bound on the impact of homeland security on private-sector productivity.

Another relevant question concerns the possible magnitude of the 1.12 percent drop in labor productivity. The most straightforward way to approach this question is to ask, at the current input levels, how much could workers assigned to homeland security have produced in measured output? In other words, suppose that there had been a 1.12 percent increase in hours used for productive purposes; by what amount would business-sector output have increased? Business-sector output in 2001 was about $8,600 billion. Increasing the labor input by 1.12 percent would have added about $70 billion to this amount—about twice as much as the federal government plans to spend on homeland security in fiscal year 2003.

Two caveats should be observed when considering the above analysis. First, the growth-accounting exercise performed here assumes that the relative productivity and prices of the various inputs remain the same. If the homeland security efforts have a large effect on the economy, they would likely affect prices as well, and our analysis would be less germane. However, because the results suggest that the effect will be rather small, it is unlikely that prices will change so drastically as to affect the results significantly.

Second, the 2001 Aviation and Transportation Security Act will shift a large part of the responsibility for airport security from the private to the public sector. This shift might lead to an increase in private-sector productivity because it will remove a large portion of unproductive overhead costs from the private sector's payrolls and capital expenditures. The efforts of companies to allow the FBI to oversee some of their employee screening (*Wall Street Journal* 2002) would have a similar effect.

Overall, homeland security's estimated effects on productivity should be rather small. In fact, because our results are likely to overestimate these effects, the actual effects on private-sector productivity will most likely be even smaller.

Total Direct Costs of Homeland Security

Our estimates of homeland security spending in the public sector combined with estimated costs in the private sector suggest that the total annual direct costs of these efforts will amount to $72 billion, or 0.66 percent of 2003 GDP (Table 3).

Table 3
Total Annual Direct Costs of Homeland Security

Scenario	Cost (Billions of Dollars)	Cost (Percentage of 2003 GDP)[a]
Federal homeland security budget	38.0	0.35
Additional homeland security spending by state and local governments	1.3	0.01
Doubling of security-related labor inputs	25.0	0.23
Doubling of security-related capital inputs	7.8[b]	0.07
Total direct costs	72.1	0.66

Source: Author's calculations.

[a]Based on Congressional Budget Office (2002) estimates of 2003 GDP.
[b]Based on an amortization of the $102 billion 2001 replacement value of electronic security systems at a 6 percent interest rate for twenty years.

Indirect Effects of Homeland Security Measures

Although the homeland security efforts of the public and private sectors could have many indirect effects on the economy, evidence of these effects thus far has been somewhat speculative. Here, we discuss two effects that have received much attention in recent months: the possibility that the tightening of airport security will increase waiting times at airports, and the fear that homeland security will draw resources away from research and development efforts and lower the rate of technological change, in turn reducing the outlook for long-run productivity growth.

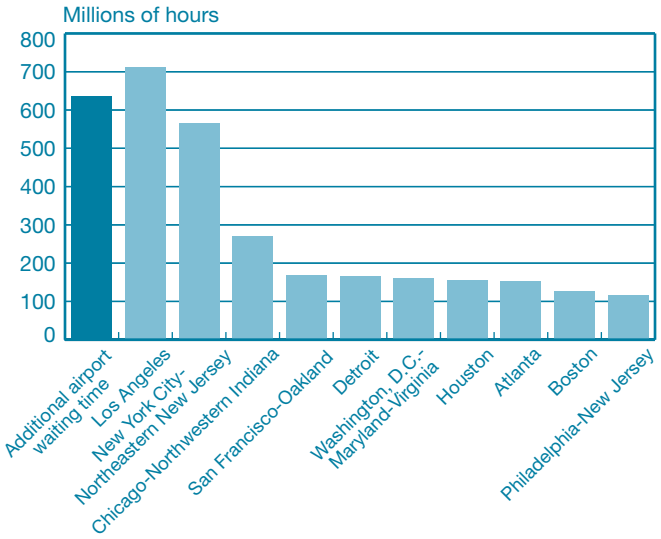
Aviation Security Effects

Part of the proposed homeland security budget, $4.8 billion to be precise, is to be spent on the improvement of aviation security. The 2001 Aviation and Transportation Security Act led to the installment of a new Transportation Security Administration, which will be responsible for the screening of passengers and baggage. At the heart of the legislation is the federalization of airport security, which involves the federal government hiring 30,000 airport security workers to perform these screenings. The law also provides funding for the purchase of equipment necessary to screen all checked baggage for explosives as well as for strengthening cockpit doors and placing more air marshals on flights.

All of these costs are already accounted for in the direct cost estimates of public-sector expenditures. However, there is a major concern that tightening security at airports will significantly increase the amount of passenger waiting time before flight boarding. (Chart 4 compares the loss in hours due to increased airport security for the ten regions with the largest time loss due to traffic congestion.) Navarro and Spencer (2001), for example, argue that the bulk of the costs of increased aviation security—$8 billion to $32 billion annually—will be attributable to this additional waiting time. Their calculation is based on the assumption that about 550 million passengers will spend an additional ninety minutes in the airport before boarding their flights and that their time is worth between $10 and $40 an hour.

However, by using more detailed data, we obtain an estimate of the cost of increased airport delays that is on the lower end of the range reported by Navarro and Spencer. First, the estimate that increased security will lead to a ninety-minute delay per passenger probably exaggerates the time loss. Poole (2001) has already observed that the new security standards

Chart 4
Total Hours Lost Due to Increased Airport Delays and Road Congestion



Sources: Air Transportation Association of America, 2001 annual report; Texas Transportation Institute, Urban Mobility Study (2001).

Note: Data are for 1999. "Additional airport waiting time" is assumed to be one hour per 1999 revenue passenger.

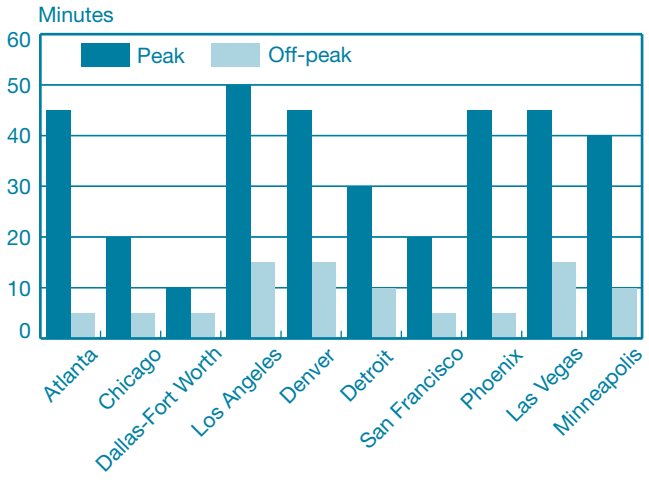
imposed are not much higher than those in major international airports around the world. This observation suggests that airport waiting times in the United States would become similar to those in other countries. Moreover, data from Delta Airlines on expected waiting times for curbside check-in, ticketing, and security checks at ten major U.S. airports indicate that none of these waiting times exceed sixty minutes during peak hours and all of them were fifteen minutes or less during off-peak hours (Chart 5). Hence, it seems reasonable to assume that the average passenger faces an extra hour of waiting time, rather than ninety minutes, due to increased security standards. This will be the scenario upon which our calculation in this section is based. Note that, given the data in Chart 5, this still seems a rather high increase, suggesting that the estimate presented here can again be considered an upper bound.

There are two types of flyers. The first are business travelers; the time they spend at the airport is measured as a labor input and thus an increase in their waiting time affects productivity. The second are leisure travelers; they fly on their own time, which is not accounted for as a productive input.

Consider business travelers. Based on data from the Travel Industry Association of America (2000), one can estimate that of the 636 million passengers boarding planes in 1999, a bit

Chart 5

Estimated Security-Check Waiting Times at Ten Major U.S. Airports



Source: Delta Airlines.

Note: Data are for May 16, 2002.

more than one-third were business travelers. The same data suggest that the average annual household income of these business travelers was $76,100, compared with a mean of $55,000 for the U.S. population. Therefore, airport delays would have led to a loss of about 215 million productive hours. If we assume that all of these business travelers were employed in the nonfarm business sector, there would have been a loss of 0.09 percent of the total hours input and a loss of 0.12 percent of the wage-adjusted labor input for the nonfarm business sector in 1999. The productivity effect of such a loss appears in Table 2.

Now, suppose that leisure travelers will also spend an additional hour waiting. Their waiting time, however, is not a productive input and, as such, does not affect productivity. In its calculation of the cost of traffic congestion, the Texas Transportation Institute (2000) valued an hour in 1999 to equal $12.40. Thus, if we were to value the hours lost by leisure travelers due to an additional hour of waiting time in 1999, their lost time would be worth $5.3 billion. Add to this amount the value of time lost by business travelers, about $6.4 billion, and the total value of this time lost would have been $11.8 billion. To put this figure in perspective, it is slightly lower than the $12.5 billion regional cost of congestion in the Los Angeles area, also estimated by the Texas Transportation Institute. This $11.8 billion in lost time is also at the lower end of the range of $8 billion to $32 billion reported by Navarro and Spencer (2001).

Furthermore, there are two other reasons why these results might overestimate the cost of delays. First, unlike time spent in traffic, time spent waiting at airports can be used for other, productive activities, such as shopping, dining, and working. The latter is especially true for business travelers, who can work using cell phones and laptops. Second, if business travel becomes a huge inconvenience, firms will continue to reduce business trips, as they have been doing since September 11. Although such a substitution effect would reduce the amount of time spent waiting in airports, it would also have a negative effect on output by reducing demand for many of the services used by business travelers.

Long-Run Productivity Effects

The above analysis suggests that the homeland security efforts will have a small effect on the *level* of productivity. However, could they also affect our outlook for the *growth* of productivity? This possibility was raised immediately after September 11, when *BusinessWeek* (2001) suggested that the homeland security efforts might indeed have a long-run effect on economic growth.

Evidence from many recent studies, such as Jorgenson and Stiroh (2000) and Oliner and Sichel (2000), suggests that the engine of the growth spurt of the economy in the late 1990s was the unprecedented investment levels, especially in computers and related equipment. The innovations in computer technology underlying these investments are driven by research and development programs. Hence, if the homeland security efforts were to reduce the R&D expenditures associated with these technologies, the productivity growth rate might be lowered.

The proposed research and development expenditures are not likely to be of such a magnitude as to impede private-sector R&D initiatives.

There are two main channels through which private-sector R&D might be affected by the homeland security efforts. First, homeland-security-related R&D might displace R&D expenditures on productivity-enhancing technologies. Moreover, if the demand for homeland-security-related equipment increases, it might lead to a shift in the private R&D

portfolio from innovations that improve output-producing technologies to security-related technologies that do not produce any measurable output. Second, a decrease in overall productivity would lower the returns to the innovations obtained from the R&D efforts.

We begin by discussing the second channel. Because our analysis suggests that the overall productivity effect of the various homeland-security-related programs will likely be fairly low, homeland security does not seem to be a deterrent to future R&D efforts. Hence, this channel will probably not play a significant role in influencing private-sector R&D.

Although it is not easy to gauge the magnitude of the first channel, we can make a reasonable estimate. We do not know how much of the private-sector R&D expenditures will be shifted from improving computers and other inputs that yield measurable output to improving security systems, yet we do know what the U.S. government proposes to spend on homeland security R&D. The 2003 budget includes "an aggressive $2.4 billion research and development program to develop technologies that will strengthen our bioterrorism response capabilities" (Office of Management and Budget 2002).

This proposed $2.4 billion in additional spending is relatively small compared with total R&D spending in the economy. In 2000, the U.S. public and private sectors spent about $265 billion on research and development; defense R&D was $24 billion, or about 9 percent of the total. If we added the currently proposed $2.4 billion of bioterrorism response R&D to this $24 billion, we would find that defense R&D spending would have been only 10 percent of total R&D spending in 2000. This percentage would be one of the lowest shares reported since the National Science Foundation began collecting data on defense R&D expenditures in 1972. Hence, the proposed research and development expenditures are not likely to be of such a magnitude as to impede private-sector R&D initiatives.[11]

Conclusion

Our evidence suggests that the economic costs of homeland security will be relatively small, and that they are unlikely to have major effects on the fiscal discipline of the government or on productivity in the private sector. Proposed government spending on homeland security is expected to account for about 0.35 percent of GDP in 2003—an amount only one-tenth the size of national defense outlays. In conjunction with this spending, even if the private sector were to double its security-related inputs, we estimate that the total annual direct costs of homeland security would be only $72 billion, or 0.66 percent of 2003 GDP. Moreover, such a doubling of inputs would at most reduce the private sector's labor productivity level by 1.12 percent.

We attach two caveats to our conclusion. First, our results do not suggest that the damage of the September 11 terrorist attacks is negligible; the findings focus solely on the economic effects of the expenditures undertaken to prevent and prepare for future incidents. Second, the results do not suggest that homeland security is unimportant. Our study is essentially only the cost side of a full cost-benefit analysis. The benefits of homeland security are, unfortunately, not always easy to measure: one simply cannot observe how many terrorist activities have been prevented because of increased security.

Clearly, it is difficult to put a value on the heightened sense of safety that the homeland security program provides. Nevertheless, given its relatively small expenses, even if the program prevented just one major incident over the next few years, the return on homeland security expenditures would be high.

Endnotes

1. The most comprehensive study of federal funding of anti-terrorism programs before 1998 was conducted by the U.S. General Accounting Office (1997).

2. On October 8, 2001, the Office of Homeland Security was established by executive order to develop and coordinate the federal homeland security program. The National Homeland Security and Combating Terrorism Act, introduced on May 2, 2002, calls for the establishment of a formal Department of National Homeland Security at the cabinet level.

3. O'Hanlon et al. (2002) recommend pursuing a slightly broader agenda that would increase federal spending on homeland security to $45 billion to $50 billion annually.

4. The U.S. Conference of Mayors notes that this figure is still preliminary, pointing out that it is a revised estimate, published in January 2002, of an October 2001 estimate of $1.5 billion for 2002 plus the last three months of 2001. This revision was necessary because the January survey suggested that cities had already spent more than $500 million on additional homeland security measures from September 11, 2001, through January 1, 2002.

5. See Office of Management and Budget of the City of New York (2002).

6. One might also ask how U.S. expenditures on combating terrorism compare with similar expenditures of other industrialized countries. Unfortunately, this question is difficult to answer because—as was the case in the United States prior to 1998—most countries do not account for their counter-terrorism expenditures separately.

7. O'Hanlon et al. (2002) propose that the Bureau of Economic Analysis account for security-related expenditures in its National Income and Product Accounts.

8. We use the terms "private" and "business" interchangeably because the data do not allow us to distinguish between private and public enterprises.

9. To consider a tripling of these inputs—assuming that the federal government will roughly triple the share of GDP devoted to homeland security—one would simply multiply the results presented here by 1.5.

10. Here, we amortize the $102 billion of capital expenditures involved in doubling the number of electronic security systems. Amortization at a 6 percent interest rate over twenty years yields an annual expense of $7.8 billion.

11. In principle, one can identify many other possible indirect effects of homeland security on the economy. Unfortunately, quantification of these effects is difficult and at best speculative.

One concern is that firms' increased uncertainty about their supply and distribution channels, as well as their demand, would induce them to target higher inventory levels. Preliminary evidence on inventory investments after September 2001 suggests that this reaction has not occurred (U.S. Department of Commerce 2002). Another concern is that increased trade costs might hamper trade and growth (see, for example, World Bank [2001] and Organisation for Economic Co-Operation and Development [2001]). Although anecdotal evidence suggests that trade costs have increased in the aftermath of September 11, there is no conclusive evidence of whether this increase will be permanent. For example, Andrea and Smith (2002) provide evidence that immediately after September 11, the Ontario-Michigan border-crossing times increased significantly, but by December 2001, these times had returned to levels that did not impede car manufacturing by disrupting logistical channels.

References

Anderson, David A. 1999. "The Aggregate Burden of Crime." Journal of Law and Economics 42, October: 611-42.

Andrea, David, and Brett Smith. 2002. "The Canada-U.S. Border: An Automotive Case Study." Altarum Institute Center for Automotive Research Car/Altarum Working Paper no. 2002-01.

Baily, Martin N. 2001. "Economic Policy Following the Terrorist Attacks." Institute for International Economics International Economics Policy Brief no. 01-10, October.

BusinessWeek. 2001. "Rethinking the Economy." October 1, p. 28.

Congressional Budget Office. 2002. "The Budget and Economic Outlook: Fiscal Years 2003-2012." January. Washington, D.C.

Council of Economic Advisers. 2002. "Economic Report of the President." February. Washington, D.C.

Jorgenson, Dale, and Kevin Stiroh. 2000. "Raising the Speed Limit: U.S. Economic Growth in the Information Age." Brookings Papers on Economic Activity, no. 1: 125-211.

Navarro, Peter, and Aron Spencer. 2001. "September 11, 2001: Assessing the Costs of Terrorism." Milken Institute Review, fourth quarter: 16-31.

New York Times. 2001. "Cities and States Say Confusion and Cost Hamper U.S. Security Drive." December 10, p. B1.

Office of Management and Budget. 2001. "Annual Report to Congress on Combating Terrorism." Washington, D.C.

———. 2002. "Budget of the United States Government, Fiscal Year 2003." Washington, D.C.

Office of Management and Budget of the City of New York. 2002. "Analysis of Agency Budgets: Mayoral Agencies." The City of New York Executive Budget for Fiscal Year 2003.

O'*Hanlon, Michael, Peter Orszag, Ivo Daalder, I. M. Destler, David Gunter, Robert Litan, and James Steinberg*. 2002. Protecting the American Homeland: A Preliminary Analysis. Washington, D.C.: Brookings Institution.

Oliner, Stephen, and Daniel Sichel. 2000. "The Resurgence of Growth in the Late 1990s: Is Information Technology the Story?" Journal of Economic Perspectives 14, no. 4 (fall): 3-22.

Organisation for Economic Co-Operation and Development. 2001. Economic Outlook, no. 70.

Poole, Robert W. 2001. "Learn from Experience on Airport Security." The Heritage Foundation Backgrounder, October 15.

RBC Capital Markets. 2001. "Safe and Sound: A Treatise on Internet Security."

Stiroh, Kevin. 2001. "What Drives Productivity Growth?" Federal Reserve Bank of New York Economic Policy Review 7, no. 1 (March): 37-59.

Texas Transportation Institute. 2000. "Urban Mobility Survey."

Travel Industry Association of America. 2000. "1999 Survey of Business Travelers."

U.S. Department of Commerce, Bureau of Economic Analysis. 2002. "National Income and Product Accounts."

U.S. Department of Labor, Bureau of Labor Statistics. 1997. BLS Handbook of Methods. Washington, D.C.

U.S. General Accounting Office. 1997. "Combating Terrorism: Spending on Governmentwide Programs Requires Better Management and Coordination." GAO/NSIAD-98-39. Washington, D.C.

Wall Street Journal. 2002. "Companies Want the FBI to Screen Employees for Suspected Terrorists." February 6.

World Bank. 2001. "Poverty to Rise in Wake of Terrorist Attacks in U.S." News Release no. 2002/093/S.

Michael J. Fleming and Kenneth D. Garbade

When the Back Office Moved to the Front Burner: Settlement Fails in the Treasury Market after 9/11

- Following the September 11 attacks, many sellers of Treasury securities failed to meet their obligation to deliver the securities on the scheduled date. Settlement "fails" jumped from $1.7 billion a day in the week ending September 5 to $190 billion a day in the week ending September 19.
- Fails rose initially because of the destruction of trade records and communication facilities. They remained high because the method typically used to avert or remedy a fail—borrowing a security through a special collateral repurchase agreement—proved as costly as failing to deliver the security.
- The U.S. Treasury responded to the fails problem by reopening the on-the-run ten-year note. The increased supply made borrowing the note more attractive than failing.
- Alternative solutions to chronic fails include the creation of a Treasury facility that could lend specific securities on a temporary basis and the institution of a penalty fee for fails.

On Thursday, October 4, 2001, the U.S. Treasury announced an unprecedented "snap," or same-day, auction of a coupon-bearing security. The auction reopened the on-the-run ten-year note and increased the outstanding supply of the note from $12 billion to $18 billion. The Treasury stated that the offering had "nothing to do with an increase of funding needs on our part" (*Bloomberg* 2001a). Rather, it sold the securities to help resolve an extraordinary volume of settlement fails precipitated by the attacks of September 11.

A settlement fail occurs when securities are not delivered and paid for on the date originally scheduled by a buyer and seller.[1] Fails are important because they expose market participants to the risk of loss in the event of counterparty insolvency. The prospect of such loss leads participants to devote resources to monitoring and controlling counterparty exposure and could, in an extreme case, lead them to limit their secondary-market trading. Treasury's statement that it reopened the ten-year note to avoid damage to "the price discovery process and the smooth operating of the Treasury market" (*New York Times* 2001) suggests that policymakers were aware of the latter possibility.

This article describes the institutional and economic setting of the fails problem and suggests why that problem led

Michael J. Fleming is a research officer and Kenneth D. Garbade a vice president at the Federal Reserve Bank of New York.

The authors thank Kenneth Abbott, Peter Bakstansky, April Bang, Anne Baum, David Buckmaster, Tom Daula, Mary Ann Davis, Robert Elsasser, Eric Foster, Kenneth Guentner, David Harrison, Ivan Hurwitz, Jeff Huther, Anthony Isola, John Kambhu, Jim Mahoney, Frank O'Connor, Carol Osler, Omer Oztan, Deborah Perelmuter, Lawrence Radecki, Kimberly Reese, Raymond Stancil, Robert Toomey, Hank Wiener, two anonymous referees, seminar participants at the Federal Reserve Bank of New York, and especially Allen Clark and Jeffrey Ingber for assistance and comments. The views expressed are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System.

policymakers to depart so significantly from previous debt management practices. The next section sets the stage by reviewing how investors establish beneficial ownership of Treasury securities and how those securities are traded in the secondary market. We then describe the machinery that supports the secondary market: the settlement process and repurchase agreements (RPs) for financing long and short positions. A description of how settlement fails come about in the normal course of trading is then offered, along with an explanation of how fails are cured. Next, we review conditions in the Treasury market following September 11, explain how those conditions led to an extraordinary volume of fails, and describe the responses of the Federal Reserve and the Treasury. Alternative mechanisms for alleviating chronic fails are then examined.

Owning and Trading Treasury Securities

Most marketable Treasury securities are owned, directly or indirectly, through the Fedwire Book-Entry Securities Transfer System (Fedwire), an electronic safekeeping and transfer system operated by the Federal Reserve and the Treasury (Stigum 1988, Chapter 7). Only depository institutions and certain other designated entities can open Fedwire accounts and own Treasury securities directly in their own accounts. Other investors establish ownership indirectly through custodial relationships with depository institutions in which a custodian institution holds investors' commingled securities in its Fedwire account and maintains a record of the securities' beneficial owners. In some cases, there may be one or more additional custodians standing between a depository institution and the beneficial owners (Martin 1985).

Most secondary-market transactions in Treasury securities are executed with dealers that make markets in the securities (Fleming 1997; Dupont and Sack 1999). Treasury dealers range from regional banks and small broker-dealer firms to large, nationally prominent banks and broker-dealers. A "primary" dealer is one that has agreed to make markets to the Federal Reserve Bank of New York when it is conducting open market operations and that has satisfied other criteria prescribed by the New York Fed.[2]

Treasury dealers also trade among themselves to manage their inventory positions and risk exposures. They sell securities in the course of reducing positions acquired from customers, they buy securities to rebuild inventories depleted by sales to customers, and they purchase and sell highly liquid on-the-run issues to manage their exposure to interest rate risk. (An on-the-run security is the most recently auctioned security in a particular series, for example, the most recently auctioned ten-year note. An off-the-run security is one that is no longer on the run. Fleming [2002] documents liquidity differences between on-the-run and off-the-run securities.)

Most marketable Treasury securities are owned, directly or indirectly, through the Fedwire Book-Entry Securities Transfer System.

Most inter-dealer trading is conducted through brokers that sponsor electronic trading systems where dealers can enter bids and offers, hit a bid posted by another dealer, and lift the offer of another dealer (Boni and Leach 2001). Trading through an inter-dealer broker is on a blind basis: the broker does not disclose the name of the buyer to the seller, or the name of the seller to the buyer. During 2001, primary dealers traded an average of $306 billion a day in Treasuries: $150 billion through brokers and $156 billion otherwise.[3]

Settling Secondary-Market Transactions

To understand the nature of the settlement problem in the Treasury market after September 11, one needs to understand how Treasury securities settle in normal times. Most secondary-market transactions are for settlement, that is, delivery of securities to the buyer and payment of the invoice price to the seller, one business day after the trade date. This is known as "regular" settlement. However, a buyer and seller may agree to some alternative settlement, including "cash" settlement on the trade date or "forward" settlement on a date following the next business day.

Dealers play a central role in the settlement process because the vast majority of secondary-market transactions are either between an investor and a dealer or between two dealers. Many of the largest dealers maintain custodial accounts at one of two "clearing" banks—JPMorgan Chase Bank (JPMC) and the Bank of New York (BoNY)—that offer specialized dealer settlement services.[4]

Bilateral Settlement

Suppose an investor sells Treasury notes to a dealer for regular settlement. Following negotiation of the terms of the sale (and usually after the close of business on the trade date), the investor instructs its custodian depository institution to deliver the notes and to collect the invoice price on the settlement date.

On the settlement date, the investor's custodian instructs Fedwire to transfer notes that the custodian knows to be beneficially owned by the investor from the Fedwire account of the custodian to the Fedwire account of the dealer's clearing bank, to collect the invoice price from the Fedwire account of the dealer's clearing bank, and to notify the dealer's clearing bank that the notes are to be credited to the dealer's custodial account. This is known as "delivery versus payment" because delivery of the securities occurs simultaneously with collection of the invoice price (Bank for International Settlements 1992). The funds are credited directly to the Fedwire account of the investor's custodian, and ultimately to the investor's custodial account.

A similar process takes place if an investor purchases notes from a dealer. On the settlement date, the dealer instructs (through software provided by its clearing bank) Fedwire to transfer notes beneficially owned by the dealer from the Fedwire account of the dealer's clearing bank to the Fedwire account of the investor's custodian and to collect the invoice price.

If the dealer's clearing bank and the investor's custodian are the same depository institution, securities are transferred from the seller to the buyer, and funds are transferred from the buyer to the seller, on the books of the common custodian and without Fedwire activity.

A sale of less than $50 million (principal amount) is usually delivered in a single block. Larger sales are delivered in multiple separate blocks of $50 million each plus (if required) a "tail" piece.[5] This means that a dealer that sells $175 million of two-year notes to a single buyer does not have to accumulate (during the course of the settlement day) a $175 million position in the notes before making delivery. The dealer can deliver part of the sale as soon as it has assembled a $50 million block, and deliver more later as it receives notes from its own purchases.

The Government Securities Clearing Corporation

Prior to 1989, bilateral settlement of inter-dealer trades compelled dealers to redeliver arriving securities repeatedly. The Government Securities Clearing Corporation (GSCC) was organized in the late 1980s to simplify the settlement process, to reduce the volume of Fedwire transfers, and to mitigate risk arising from counterparty exposures. Until 2002, GSCC was owned primarily by its members and operated on a not-for-profit basis. (In January 2002, GSCC became a wholly-owned subsidiary of Depository Trust & Clearing Corporation.) GSCC's membership includes—but is not limited to—all of the primary dealers, a number of banks and broker-dealers that are not primary dealers, and the inter-dealer brokers.[6]

The Government Securities Clearing Corporation facilitates the comparison of trades between members, reduces deliveries to the smallest possible volume, and steps in as the counterparty in the net settlement of all compared trades.

The Government Securities Clearing Corporation facilitates the comparison of trades between members, reduces deliveries to the smallest possible volume, and steps in as the counterparty in the net settlement of all compared trades.

GSCC is not a depository institution and does not have a Fedwire account. It divides its settlement business between BoNY (for notes) and JPMC (for bills, bonds, STRIPS, and inflation-indexed securities).

Trade Comparison

A trade is said to be "compared" when both the buyer and the seller have acknowledged the transaction and agreed to its terms, usually to a third party. Following a direct trade between two GSCC members, each member sends an advisory message to GSCC identifying the counterparty, the security, the quantity of the security, the invoice price, and the settlement date. GSCC deems the trade compared when it has matching information from the two parties.

Following a brokered trade between two GSCC members, the broker sends two messages to GSCC: one identifying the buyer and the terms of the trade, the other identifying the seller and the terms of the trade. Additionally, the buyer and seller each send messages of their own, identifying the terms of the trade, the broker, and their role (buyer or seller) in the transaction. GSCC compares the seller's message with the corresponding broker's message and the buyer's message with

the other broker's message separately. This two-part process is required because brokered trading is on a blind basis: a buyer and seller do not know each other's identity either before or after agreeing to a trade.

Net Settlement

After the cessation of trading at the end of a business day, GSCC computes, from all compared trades, the net obligation of each member to either receive or deliver securities the next business day, on an issue-by-issue basis.[7] Then, in a process known as "novation," that is, the substitution of a new legal obligation for an old one, GSCC steps in as the counterparty to every member. A member with a net obligation to deliver a security is instructed to deliver the security to GSCC; a member with a net obligation to receive a security is advised that it will receive the security from GSCC. (The deliveries to and receives from GSCC net to zero in every security. GSCC does not trade securities for its own account and is only a conduit for settlement purposes.) At the same time, GSCC instructs the clearing banks what securities it expects to receive and how much to pay for the securities, and it instructs the banks where the securities should be redelivered and how much to collect upon redelivery.

The primary advantage of this net settlement system is that each GSCC member need only deliver its net sales (or receive its net purchases), instead of both delivering its gross sales and receiving its gross purchases. On an average day in 2001, GSCC compared and netted about 55,000 transactions with a market value of about $1.4 trillion. The transactions netted down to an average of about 11,000 deliveries with a value of about $400 billion (Government Securities Clearing Corporation 2002, p. 2).

GSCC creates deliver and receive instructions in pairs. Each instruction to deliver a security to GSCC is paired with, or "bound" to, an instruction to receive an equal size block of the same security from GSCC. After the opening of Fedwire on a settlement day, GSCC begins to receive blocks of securities pursuant to the delivery instructions that it created the preceding night and promptly (in a matter of seconds) redelivers those blocks pursuant to its own novated delivery obligations. GSCC's practice of creating deliver and receive instructions in matched and bound pairs facilitates immediate redelivery because it never has to wait for additional securities to come in to make up an outgoing block.

Transaction Adjustment Payments

All deliveries of a security to GSCC, and all redeliveries of the same security from GSCC, are made against payment at a common current market value (CMV). The CMV of a security is a representative price and is not necessarily equal to the price of any particular transaction. GSCC marks the transactions of each of its members in a security to the security's CMV by adding up the net excess of a member's purchase prices over the security's CMV and subtracting the net excess of the member's sale prices over the CMV. These net amounts are aggregated across all of the member's securities to produce a transaction adjustment payment. If the adjustment payment is greater than zero, the member is required to pay it into a GSCC account at BoNY by 10 a.m. on the settlement date. If it is negative, GSCC agrees to pay it out to the member by 11 a.m.

Financing Long and Short Positions

Some market participants, especially dealers and hedge funds, commonly commit themselves to paying amounts in excess of their cash balances when they purchase Treasury securities. Moreover, they sometimes sell more securities than they own, or sell securities short. In the former case, they have to borrow money to make the payment due upon delivery. In the latter case, they have to borrow securities to make delivery and get paid. Borrowing money to pay for a purchase, and borrowing securities to settle a short sale, are "financing" transactions, which are usually done with repurchase agreements. Repurchase agreements are important to understanding the fails problem after September 11 because they can be used to avoid or cure settlement fails and because they themselves may fail to settle.

To borrow money with an RP, a dealer sells securities (typically for cash, or same-day, settlement) and simultaneously agrees to repurchase the same securities from the buyer at a higher price on a future date. As illustrated in Box 1, an RP is tantamount to a collateralized loan. The proceeds of the sale are the principal amount of the loan, and the excess of the repurchase price over the sale price is the interest paid on the loan.

The opposite occurs when a dealer has to borrow securities: the dealer buys (again, usually for cash settlement) the securities that it needs from an investor and simultaneously agrees to sell the same securities back to the investor at a higher price on a future date. This transaction is an RP from the investor's perspective: a sale of securities today, coupled with an agreement to repurchase the securities on a future date. The dealer is said to enter into a "reverse repurchase agreement"—

"reversing in" the securities that it needs against lending money, or borrowing securities using cash as collateral.

Corporate and municipal treasurers and money market mutual funds have money to lend for short periods of time and are important sources of funds for dealers. Other investors, such as pension funds and life insurance companies, own Treasury securities and are in a position to lend them to dealers. However, one might wonder why a pension fund or life insurance company would want to borrow a dealer's money. The answer lies in the distinction between a general collateral RP and a special collateral RP.

Box 1
A Repurchase Agreement

On Monday, December 1, dealer A wants to finance $10 million principal amount of Treasury notes. The notes have a 5 percent coupon payable in semi-annual installments on May 15 and November 15 and are quoted at a price of 99 percent of principal. Investor B agrees to lend the dealer money on an overnight repurchase agreement at a rate of 3 percent per annum.

Assume for simplicity that the dealer sells the notes to the investor at the quoted price plus accrued interest for cash settlement on December 1. The accrued interest on the notes is computed using the 5 percent coupon rate and an actual-over-actual-day count. There are 16 days from November 15 to December 1 and 182 days from November 15 to May 15 (assuming the absence of a leap day), so the accrued interest to December 1 is:

$$\text{accrued interest} = \frac{16}{182} \cdot \frac{1}{2} \cdot 5.00 = .219780 \text{ percent of principal amount.}$$

The amount borrowed, that is, the amount due upon delivery of the notes to investor B, is $9,921,978 ($9,921,978 = 99.219780 percent of $10 million).

On Tuesday, December 2, dealer A repurchases the notes for the amount borrowed on December 1 plus interest on that amount at the rate of 3 percent per annum, calculated with an actual-over-360-day count:

$$\text{repurchase amount} = \$9{,}921{,}978 + \frac{1}{360} \cdot 3 \text{ percent of } \$9{,}921{,}978 = \$9{,}922{,}805.$$

General Collateral Repurchase Agreements

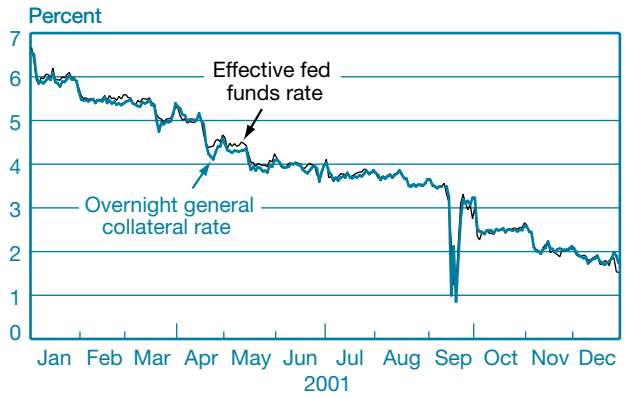
A general collateral repurchase agreement is an RP in which the lender of funds is willing to accept any of a variety of Treasury securities as collateral. The lender is concerned primarily with earning interest on its money and with having possession of securities that can be sold quickly in the event of a default by the borrower. Reflecting the relative indifference of the lender to the specific identity of the collateral, interest rates on general collateral RPs are commonly negotiated early in the business day—more than half of all general collateral RPs are negotiated before 9 a.m.—while collateral may not be assigned until late morning.

As shown in Chart 1, interest rates on overnight general collateral RPs are usually quite close to contemporaneous rates on overnight loans in the federal funds (fed funds) market. This comports with the essential character of a general collateral RP as a device for borrowing and lending money.

Special Collateral Repurchase Agreements

A special collateral repurchase agreement is an RP in which the lender of funds has identified a particular security as the only acceptable collateral. The rate on a special collateral RP is commonly called a "specials" rate. Each Treasury security has its own specials rate. This reflects the essential character of a special collateral RP as a device for borrowing and lending securities.

Chart 1
General Collateral and Effective Fed Funds Rates



Sources: Bloomberg; GovPX.

The owner of a Treasury security that a dealer wants to borrow may not have any particular interest in borrowing the dealer's money, but can nevertheless be induced to lend the security if it is offered an opportunity to borrow at a specials rate below where it can relend the same funds on a general collateral RP. For example, if the rate on a special collateral RP is 2 percent and the rate on a general collateral RP is 3 percent, then—as shown in Exhibit 1—an investor can earn a 100-basis-point spread by borrowing money on the special collateral RP and relending the money on a general collateral RP.

The Specials Rate for a Security and the "Specialness" of the Security

The difference between the general collateral RP rate and the specials rate for a security is a measure of the "specialness" of the security (Duffie 1996; Keane 1996; Jordan and Jordan 1997). If the demand to borrow a security is modest relative to the supply of the security available for lending, a borrower of the security will usually be able to lend its money at a rate no lower than about 15 to 25 basis points below the general collateral rate.

If the demand to borrow a security is strong, or if the supply of the security available for lending is limited, the specials rate for the security may be materially below the general collateral rate and the specialness spread correspondingly large. In this case, a dealer borrowing the security has to sacrifice a significant portion of the interest that it could have earned from lending its money in the general collateral market. Conversely, a holder of the security will be rewarded with a "bargain" loan rate. Charts 2 and 3 show specials rates for on-the-run five- and ten-year notes, respectively, in 2001, and Table 1 shows average specialness spreads for these and other securities during the first eight months of 2001.[8]

In cases of exceptionally strong demand, or exceptionally limited supply, the specials rate for a security can be driven to zero or nearly zero. Charts 2 and 3 show several such instances, including the period from late January to early February for the five-year note and the period from late April to early May for the ten-year note. Other instances of extremely low specials rates have been documented by Cornell and Shapiro (1989), Jordan and Jordan (1997, pp. 2058-9), and Fleming (2000, pp. 229-31).

Supply, Demand, and Equilibrium in a Specials Market

The specialness spread for a security is the economic price of, or fee for, borrowing the security and fluctuates to balance the demand for borrowing with the supply available for lending. Ceteris paribus, a larger spread elicits greater supply because it offers a greater reward to owners of the security who borrow funds in the specials market (thereby lending the security) and relend the funds in the general collateral market.

A larger spread also reduces the demand for borrowing a security. Borrowing demand stems primarily from the need to finance short positions. Ceteris paribus, a larger specialness spread increases the cost of financing a short position and thus reduces the attractiveness of being short. Additionally, as the specials rate for a security approaches zero, market participants

Exhibit 1

Lending Collateral (and Borrowing Money) on a Special Collateral Repurchase Agreement and Relending the Money on a General Collateral Repurchase Agreement

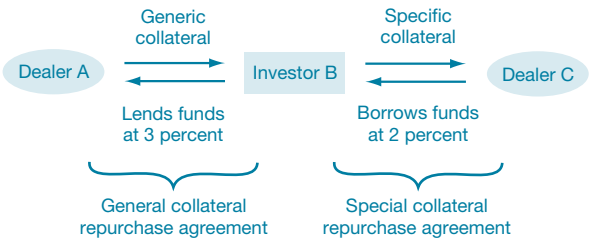
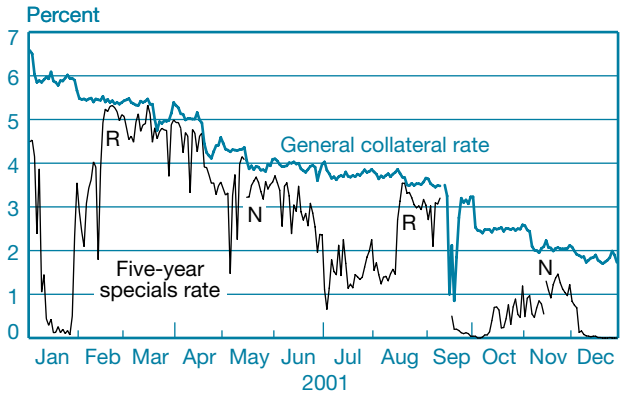


Chart 2

RP Rates for the Five-Year Treasury Note and General Collateral

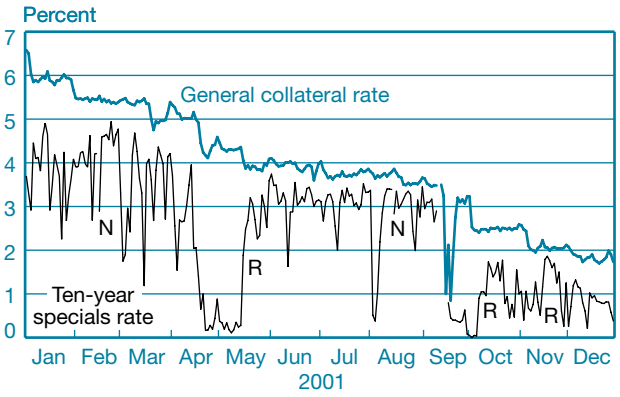


Sources: Bloomberg; GovPX.

Notes: The chart plots overnight repurchase agreement (RP) rates for the on-the-run five-year U.S. Treasury note and general collateral from January 2 to December 31, 2001. "N" indicates the issuance of a new note and "R" indicates the reopening of an existing note.

Chart 3

RP Rates for the Ten-Year Treasury Note and General Collateral



Sources: Bloomberg; GovPX.

Notes: The chart plots overnight repurchase agreement (RP) rates for the on-the-run ten-year U.S. Treasury note and general collateral from January 2 to December 31, 2001. "N" indicates the issuance of a new note and "R" indicates the reopening of an existing note.

with short positions may begin to opt to fail on their delivery obligations rather than go to the trouble of lending money at a negligible rate of interest to borrow the securities needed to make delivery. This feature is critically important to understanding the fails problem following September 11 and will be examined later in more detail.

Table 1

Specialness Spreads for Treasury Securities, January-August 2001

	On-the-Run		First-off-the-Run	
Security	Mean	Standard Deviation	Mean	Standard Deviation
Three-month bill	6.4	6.9	—	—
Six-month bill	5.4	3.2	—	—
Two-year note	107.8	101.7	82.2	92.4
Five-year note	142.2	156.2	18.5	20.0
Ten-year note	156.6	130.7	29.8	42.7
Thirty-year bond	7.9	6.9	8.0	8.4

Source: GovPX.

Notes: The table reports descriptive statistics (in basis points) for the daily differences between the overnight general collateral rate and the specials rates on the indicated U.S. Treasury securities. A first-off-the-run security is the most recently issued off-the-run security in a particular series.

Settling a Repurchase Agreement

An RP involves two settlements: one at inception and the other at maturity. Following negotiation of the terms of an agreement, and usually on the same day, the lender of securities delivers the securities to its counterparty against payment of the principal amount of the RP. On the maturity date, the borrower of securities delivers the securities back to the original owner against payment of interest plus repayment of principal.

Comparing and Settling RPs between GSCC Members

If both parties to a bilaterally negotiated RP are GSCC members, the parties can submit their agreement to GSCC's comparison and netting processes. Following negotiation of the terms of the RP, each party sends a message to GSCC identifying its counterparty and describing the terms. GSCC deems the agreement compared if it receives matching information from the two parties.

If the agreement calls for cash settlement of the starting leg, settlement of that leg is bilateral between the borrower and lender because GSCC does not net and novate transactions until after the close of business. If the agreement provides for settlement of the starting leg on a date following the trade date, settlement is included in GSCC's netting and novation process for that date. In either case, settlement of the closing leg of the RP is included in GSCC's netting and novation process.

Blind-Brokered RPs with Cash Settled Starting Legs

Inter-dealer brokers arrange RPs as well as outright transactions. Prior to 1996, a broker would "give up" the name of each side of a brokered RP that provided for cash settlement of the starting leg to the counterparty and leave the two sides to settle the starting leg bilaterally. This arrangement had the disadvantage of allowing the lender of money on a general collateral RP to gain some insight into the long positions held by its counterparty, and of allowing the lender of securities on a special collateral RP to acquire information on a short position that its counterparty needed to finance.

In mid-1996, GSCC began to facilitate a less revealing way to settle the opening leg of a blind-brokered RP. Following negotiation of a blind-brokered RP, the borrower of money delivers the collateral securities to the broker against payment of the principal amount of the loan and the broker redelivers the securities (against the same payment) to the ultimate lender. GSCC makes this possible by guaranteeing broker payment and delivery obligations on compared transactions. (The broker's role in cash settlement of the starting leg of a

blind-brokered RP cannot be eliminated as long as GSCC does not net and novate transactions during a business day. This proved to be an important element of the fails problem following September 11.)

Borrowing Securities from the Federal Reserve

The Federal Reserve's System Open Market Account (SOMA) is the largest single owner of marketable Treasury securities, holding $575 billion in principal amount on December 31, 2001, or 19.4 percent of the $2.968 trillion outstanding. The Federal Reserve allows primary dealers to borrow securities from SOMA as a "secondary and temporary source of securities . . . in order to promote smooth clearing of Treasury securities."[9] Federal Reserve efforts to mitigate settlement problems after September 11 were carried out through this securities lending program.

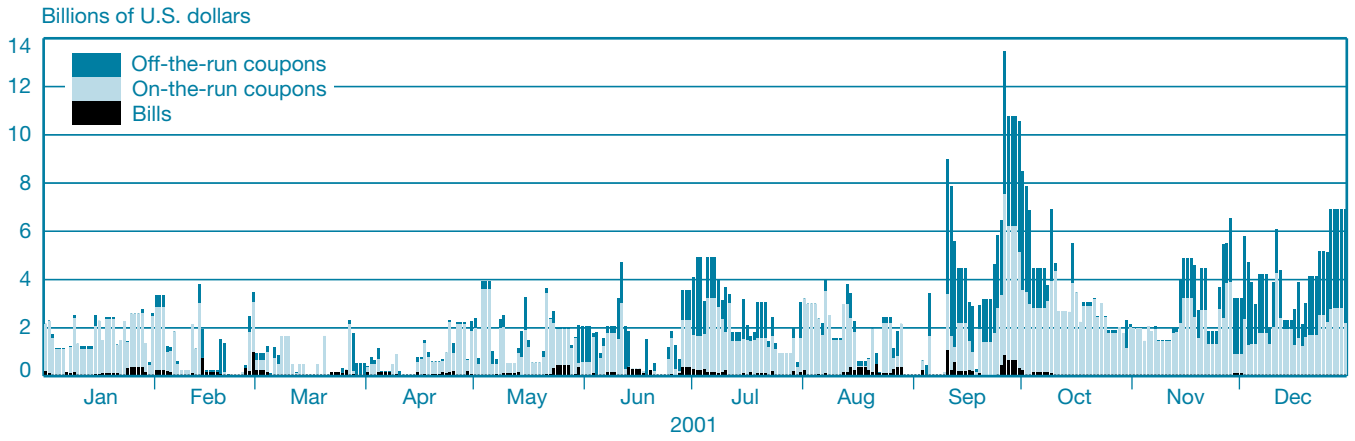
The program works as follows. Each day at noon, the New York Fed auctions a specified amount of every Treasury security in the SOMA portfolio for overnight lending. (From September 7, 1999, through September 26, 2001, the amount was 45 percent of the quantity beneficially owned by SOMA, subject to an upper limit of the amount actually in SOMA's account at the time of auction.) A participating dealer bids a borrowing fee that is economically equivalent to a specialness spread for each security that it wants to borrow. For example, a dealer might offer to pay a fee of 2 1/2 percent per annum to borrow a security if the general collateral rate is 4 percent and the specials rate for the security is 1 1/2 percent. Securities are awarded to the highest bidders at their bid rates until all bidders are satisfied or all of the available stock of a security has been allocated. Prior to September 11, a dealer had to bid a borrowing fee of at least 150 basis points and was limited to $100 million of any one security and $500 million in aggregate. Borrowings from SOMA are collateralized with other Treasury securities.

Chart 4 shows the volume of SOMA lending activity in 2001. Over the interval from January to August, primary dealers borrowed an average of $1.6 billion in principal amount of securities per day. On-the-run notes and bonds accounted for 69.7 percent of the borrowings, off-the-run notes and bonds accounted for 25.2 percent, and the balance of 5.1 percent was in bills.

Settlement Fails

A settlement fail is the failure of a transaction to settle on the date originally agreed to by the buyer and seller. The transaction can be an outright sale, or the starting or closing leg of an RP. Although the post-September 11 period was characterized by an unprecedented level of fails, fails also occur

Chart 4
SOMA Securities Lending



Source: Federal Reserve Bank of New York.

Notes: The chart plots daily lending of U.S. Treasury securities by the System Open Market Account (SOMA) of the Federal Reserve from January 2 to December 31, 2001. Total lending is broken down by lending in bills, on-the-run coupon securities, and off-the-run coupon securities.

in more normal times. During the first eight months of 2001, settlement fails in Treasury securities at primary dealers averaged $7.3 billion per day. Chart 5 shows the behavior of fails in 2001.

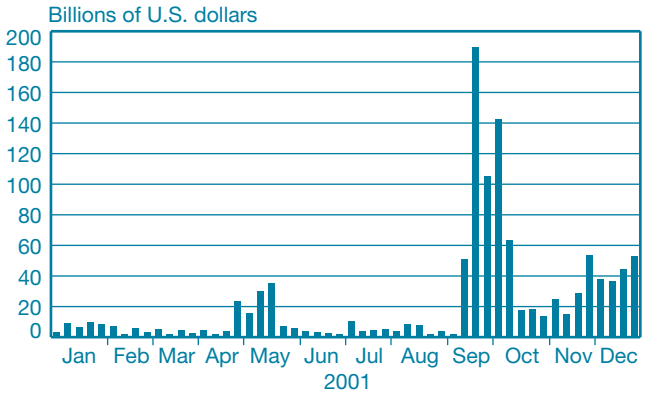
Fails occur for a variety of reasons. One source of fails is miscommunication. Despite their best efforts to agree on terms, a buyer and seller may sometimes not identify to their respective operations departments the same details for a given transaction. On the settlement date, the seller may deliver what it believes is the correct quantity of the correct security and claim what it believes is the correct payment, but the buyer will reject the delivery and reclaim its funds if it has a different understanding of the transaction. If the rejection occurs late in the day, there may not be enough time for the parties to resolve the misunderstanding.

In some cases, a seller or a seller's custodian may be unable to deliver securities because of operational problems. In an extreme example, a computer problem at BoNY on November 21, 1985, prevented the bank from issuing instructions to Fedwire to deliver securities from its custodial accounts. The bank was unable to resolve the problem until the following day. In the interim, it had to finance (at its own expense) the securities that it was unable to deliver. It borrowed in excess of $20 billion from the New York Fed and incurred interest expenses of about $5 million (*Wall Street Journal* 1985; Sender 1986).

Finally, and most commonly, a seller may be unable to deliver securities because of a failure to receive the same securities in settlement of an unrelated purchase. This can lead to a "daisy chain" of cumulatively additive fails: A's failure to deliver bonds to B causes B to fail on a sale of the same bonds to C, causing C to fail on a similar sale to D, and so on. A daisy chain becomes a "round robin" if the last participant in the chain is itself failing to the first participant.

Although the post-September 11 period was characterized by an unprecedented level of fails, fails also occur in more normal times. During the first eight months of 2001, settlement fails in Treasury securities at primary dealers averaged $7.3 billion per day.

Chart 5
Settlement Fails in U.S. Treasury Securities



Source: Federal Reserve Bank of New York.

Note: The chart plots daily average settlement fails to deliver of U.S. Treasury securities as reported by the primary dealers for the weeks ending January 3 through December 26, 2001.

The Cost of a Fail

Market participants recognize that miscues and operational problems occur from time to time, and have adopted the convention of allowing a failing seller to make delivery the next business day at an unchanged invoice price (Public Securities Association 1993, Chapter 8, Section C).[10] Settlement fails, however, are not costless.

The most important cost of a fail is that the seller loses the time value of the invoice price over the interval of the fail. (This cost may be reimbursed by the buyer if the buyer's actions caused the fail, for example, by improperly rejecting securities properly tendered by the seller, or by a third party if the third party's actions caused the fail, for example, a custodian failing to deliver securities pursuant to the instructions of the seller.) This implicit penalty (which can be quantified as the interest that could have been earned in the fed funds or general collateral markets) provides an incentive to sellers to avoid and cure fails. (There is an exactly offsetting benefit to a buyer that fails to receive securities and therefore does not have to pay for them as soon as originally scheduled: it can invest the invoice price until the securities arrive. Stigum [1988, pp. 181-5] and DeGennaro and Moser [1990] describe the relationship between fails and dealer behavior.)

Additionally, a fail exposes both the buyer and the seller to replacement cost risk. The buyer faces the risk that the seller becomes insolvent before settlement and that the price of the security increases prior to the seller's insolvency. Conversely,

the seller faces the risk that the buyer becomes insolvent and that the price of the security declines prior to the buyer's insolvency. The significance of replacement cost risk exposure may be small for a fail that does not last more than a few days, but it increases as a fail continues. Aged fails generally prompt market participants to step up their monitoring of counterparties.

A fail exposes both the buyer and the seller to replacement cost risk . . . . The significance of replacement cost risk exposure may be small for a fail that does not last more than a few days, but it increases as a fail continues.

Initially, the increased monitoring may be nothing more than a phone call to identify whether there has been a misunderstanding, but it can escalate to credit reviews and requests for ad hoc mark-to-market price adjustments.[11] Additionally, the net capital requirement for regulated brokers and dealers adopted by the Securities and Exchange Commission assesses capital charges for two types of aged fails (17 CFR 240.15c3-1(c)(2)(iv)(E) and 17 CFR 240.15c3-1(c)(2)(ix)).

Avoiding and Curing Fails

Market participants have adopted a variety of techniques to avoid fails attributable to miscommunication and to avoid and cure fails attributable to a failure to receive securities that are to be redelivered. Most importantly, GSCC's comparison process limits fails attributable to miscommunication, and its netting and novation process limits daisy chain and round robin fails among its members. For example, if member A sells $25 million of bonds to member B, B sells $25 million of the same bonds to member C, and C sells $25 million of the bonds to member D, then, after netting and novation, A owes $25 million of the bonds to GSCC and GSCC owes $25 million of the bonds to D. If A fails to GSCC (and GSCC consequently fails to D), the aggregate fails to deliver and receive will be limited to $50 million each. In the absence of netting and novation, A's failure to deliver could result in aggregate fails to deliver and receive of $75 million each.

GSCC's net settlement process does not completely eliminate round robin fails, because GSCC fails are not usually renetted every day.[12] A GSCC member that fails to deliver securities pursuant to a GSCC instruction must ultimately make that specific delivery, even if the member is a net buyer of an equal or larger amount of the same issue for settlement on the next business day. Similarly, GSCC remains obligated to redeliver the securities pursuant to its matched and bound redelivery instruction, even if the member to whom GSCC owes the securities is a net seller of an equal or larger amount of the same issue for settlement on the next business day. Thus, it is possible for a GSCC member to be failing to GSCC and for GSCC to be failing simultaneously to the same member in the same security on a delivery instruction created on a different date.

As shown in Exhibit 2, fails stemming from an inability to deliver securities because of a failure to receive the same securities can be cured (or avoided entirely) by reversing in the securities needed for delivery and delivering the borrowed securities. The reverse RP can be closed out when the seller finally makes delivery. A market participant with offsetting fails is usually better off earning the specials rate on the money it lends out against reversing in the securities needed to make delivery, even if that rate is below the general collateral rate. The alternative to reversing in securities is to fail and forgo interest altogether on the sale proceeds.

Market participants have adopted a variety of techniques to avoid fails attributable to miscommunication and to avoid and cure fails attributable to a failure to receive securities that are to be redelivered.

When Fails Can Become Chronic

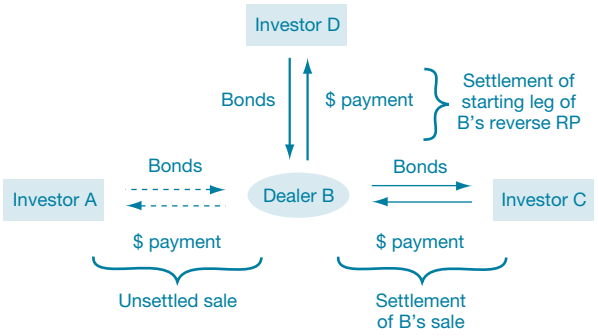
The incentive of a seller to borrow securities to avoid or cure a fail declines with the specials rate for the security. When the specials rate is near zero, a seller has little to gain lending money (at nearly no interest) to borrow the needed securities.[13] This suggests that market participants may have little incentive to break daisy chains and round robins when the specials rate for a security is near zero. This aspect of the market is important to understanding the fails problem after September 11.

A specials rate near zero becomes increasingly likely when there is strong demand for borrowing a security. As noted, demand for a security will be strong when there is substantial short interest in the security (due to hedging requirements or

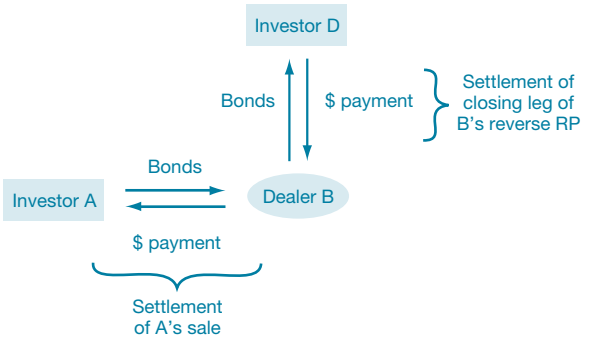
Exhibit 2

Borrowing Securities to Cure a Settlement Fail

Suppose investor A sells bonds to dealer B, who then resells the bonds to investor C. The delivery and payment obligations of A, B, and C on the morning of the settlement date are shown with dashed lines.



Assume that A fails to deliver the bonds to B, causing B to fail to C. To cure its fail to C, B can reverse in the bonds from investor D and deliver the borrowed bonds. Actual movements of bonds and funds are shown with solid lines. The continuing (unfulfilled) obligations of A to deliver bonds to B and B to pay for the bonds are shown with dashed lines.



When A finally delivers the bonds to B, B can terminate its reverse repurchase agreement (RP) with D.

expectations of interest rate changes). In addition, such short interest is usually concentrated in highly liquid on-the-run issues. This suggests why specials rates for on-the-run five- and ten-year notes fall (Charts 2 and 3) and fails rise (Chart 5) around mid-quarter refunding auctions, a time when market participants are actively engaged in taking positions based on their assessments of the relative values of different securities and the prospect for change in the level and shape of the yield curve (Garbade 1996, Chapter 8).

Additionally, the specials rate for a security will be driven to its lower limit more frequently when the fed funds rate, and hence the general collateral rate, is lower. This follows because the gross compensation earned by a lender of securities at any given specials rate is the difference between the general collateral rate and the specials rate. A lower general collateral rate rations marginal lenders of securities out of the specials market and reduces the quantity of securities available for lending at a given specials rate. This leads to a lower equilibrium specials rate and hence to a more common occurrence of the specials rate driven to nearly zero.

Settlement Fails Following September 11

The attacks of September 11 precipitated an extraordinary increase in settlement fails in the Treasury securities market. As shown in Chart 5, daily average fails jumped from $1.7 billion during the week ending Wednesday, September 5, to $190 billion during the week ending September 19. (Our fails data begin in mid-1994. The previous high-water mark had been $35 billion during the week ending May 16, 2001.) The Federal Reserve responded by relaxing restrictions on its securities lending program. When the volume of fails did not abate after three weeks and when fails began to threaten, according to Treasury, "the price discovery process and the smooth operating of the Treasury market" (*New York Times* 2001), Treasury reopened the on-the-run ten-year note.

Initial Events

American Airlines Flight 11 crashed into the north tower of the World Trade Center at 8:46 a.m. on Tuesday, September 11. United Airlines Flight 175 hit the south tower sixteen minutes later, and both towers collapsed before 10:30 a.m. The catastrophe destroyed the offices and business records of several inter-dealer brokers and damaged the facilities of Verizon Corporation, located immediately north of the north tower, thereby impairing telecommunication services throughout Lower Manhattan.

Destruction of Broker Offices and Records

Although trading stopped soon after the first attack, purchases and sales of about $80 billion of Treasury and related securities, and about $500 billion in repurchase agreements, had already been negotiated on September 11. Most of the inter-dealer

transactions had been arranged by brokers, several of which were located in the towers of the World Trade Center. Those that did not submit trade data on a real-time basis had not yet advised GSCC of the terms of many of the transactions, and they had not yet given their clearing banks instructions for redelivering securities received in settlement of the starting legs of many of the RPs. This led to two significant operational problems.

First, the starting legs of numerous blind-brokered RPs failed to settle. Under normal conditions, if dealer A financed, through broker B, a purchase or repurchase of $10 million of bonds with dealer C, dealer A would send the bonds to B's clearing bank and the broker would instruct its bank to redeliver the bonds to dealer C. (Both of these deliveries would be against payment of the principal amount of the RP.) On September 11, however, the clearing banks never received redelivery instructions for many RPs and the securities either remained in the brokers' clearing accounts or were returned to the putative borrowers.

Although trading stopped soon after the first attack, purchases and sales of about $80 billion of Treasury and related securities, and about $500 billion in repurchase agreements, had already been negotiated on September 11.

Second, GSCC could not compare numerous brokered transactions because it was missing the brokers' advisory messages. Uncompared transactions were liable to be excluded from the netting and novation process scheduled for the evening of September 11 and thrown back on GSCC's members to resolve as best they could. Resolution would, however, be extraordinarily difficult because the buyers and sellers did not know the identities of their ultimate counterparties and because the offices and records of some of the brokers had been destroyed.

Impaired Telecommunication Links

Damage to telecommunication facilities impaired connectivity between many market participants and affected BoNY especially severely (MacRae 2001; *Wall Street Journal* 2001a, d; Bond Market Association 2001). Several brokers whose offices were not destroyed were nevertheless unable to advise GSCC of trades they had arranged on the morning of September 11. Moreover, GSCC and several dealers could not verify what came into and what left their custodial accounts at BoNY, they could not advise BoNY of securities they expected to receive, and they could not give BoNY instructions for delivering securities. Additionally, GSCC was unable to verify the movement of funds into and out of its account at BoNY (GSCC Important Notice GSCC068.01).

GSCC Initiatives

GSCC took two actions on the night of September 11 to minimize the problems caused by the destruction of broker records and by the loss of telecommunication links (GSCC Important Notice GSCC070.01). First, it created, as best it could from the dealer messages that it had received, the missing broker messages on more than 2,000 trades valued at more than $70 billion.[14] This action brought the trades into GSCC's netting and novation process and centralized the reconciliation of transactions that might have otherwise proved irreconcilable. Second, GSCC assumed responsibility for starting-leg fails on almost $100 billion of blind-brokered RPs, even though it had not been able to compare many of the agreements with the brokers and their counterparties.

Consequences for Settlement Fails

GSCC classifies as a fail any delivery instruction for which it does not receive a report of good delivery (GSCC Important Notice GSCC074.01). In the absence of complete information on deliveries into and out of its account at BoNY on September 11, and as a result of its assumption of settlement fails on the starting legs of blind-brokered RPs, GSCC recorded (after the close of business on September 11) $266 billion in transactions that apparently failed to settle.[15]

After the close of business, GSCC followed its normal practice and paired off starting-leg fails on overnight RPs with the closing legs of the same agreements and canceled both legs.[16] (It marked its remaining fails to market and created new delivery instructions for those fails.[17]) Also after the close of business on September 11, GSCC netted member transactions scheduled to settle on September 12 (including the brokered transactions for regular settlement for which it had created missing broker messages), novated the net purchase or sale obligations of each of its members, and created delivery instructions for the novated obligations.

Continuing connectivity problems prevented GSCC from giving BoNY delivery instructions after the close of business on September 11 and prevented it from acquiring information on activity in its account at BoNY during the day on September 12. Consequently, GSCC recorded $440 billion in settlement fails as of the close of business on September 12.

As market participants gradually reacquired connectivity in succeeding days, they began to reconcile their incomplete and erroneous trade and settlement information. Settlement fails at GSCC dropped to $403 billion on September 13, to $225 billion on September 14, and to about $90 billion by the end of the following week.

Initial Federal Reserve Actions

The Federal Reserve reacted promptly to the September 11 attacks. The best-known dimensions of the response are the expansion of discount window loans and RPs with primary dealers that ensured bank liquidity and the integrity of the payments system (Federal Reserve Bank of New York 2002; McAndrews and Potter 2002). On the securities side, the New York Fed acted to make Treasury collateral more readily available to primary dealers. On September 11, it suspended the $500 million limit on borrowings of SOMA securities by a single dealer, as well as the issue limit of $100 million for a single dealer. (The post-September 11 changes to the securities lending program are posted at <http:// www.newyorkfed.org/pihome/news/opnmktops/2001/omo010911.html>.) As shown in Chart 4, securities borrowings increased from $100 million on September 10 to $8.9 billion on September 11 and then declined gradually in the following days, to $7.8 billion on September 12, $5.6 billion on September 13, and steadily lower amounts the following week.

Persistence of Settlement Fails

Despite painstaking efforts by market participants to reconcile their transactions, and despite the expanded volume of securities borrowings from the System Open Market Account, settlement fails in the Treasury market did not abate quickly to pre-attack levels. As shown in Chart 5, primary dealers reported daily average fails of $190 billion for the week ending September 19, $105 billion for the week ending September 26, and $142 billion for the week ending October 3.

The persistence of high levels of fails can be most immediately attributed to specials rates at or close to zero in the weeks following September 11 (Charts 2 and 3). With specials rates so low, market participants had limited incentive to reverse in securities to break the daisy chains and round robins that had appeared during the chaos of September 11 and the days immediately following. Additionally, the low specials rates may have given some market participants an incentive to fail "strategically," or to agree to lend securities on special collateral RPs and then to fail intentionally on the starting legs of the agreements (Box 2).

Specials rates were driven down to negligible levels by contractions in the supply schedules of securities available for lending. This resulted in part from the reduction in the target fed funds rate from 3 1/2 percent to 3 percent on Monday, September 17. The action reduced the gross economic return to lending Treasury securities at any given specials rate and rationed some marginal lenders of securities out of the specials market.

Box 2
Strategic Fails

When the specials rate for a security is close to zero, a market participant with no position in the security may sometimes lend the security on a term repurchase agreement (RP) and then fail intentionally on the RP's starting leg.

Suppose, for example, the three-week specials rate for a five-year note is 10 basis points and that XYZ Co. believes the two-week specials rate will be 50 basis points in one week. If XYZ contracts (in the specials market) to borrow $50 million for three weeks against lending the note, it will owe interest of $2,916.67 at the end of three weeks ($2,916.67 = (21/360) · .0010 · $50,000,000). It will owe this amount even if it fails to deliver the note during the entire three-week interval. XYZ Co. has effectively purchased an implicit option (for $2,916.67, payable in three weeks) on an exchange of $50 million for the five-year note at any time during the next three weeks for the balance of the three-week interval.

XYZ Co. could choose to let its implicit option expire unexercised and simply pay the $2,916.67 premium at the end of three weeks. However, if XYZ Co.'s expectations prove correct, it can exercise the option after one week by reversing in the five-year note for two weeks against lending $50 million at 50 basis points (earning interest of $9,722.22) and delivering the note in (delayed) settlement of its earlier negotiated three-week RP ($9,722.22 = (14/360) · .0050 · $50,000,000). The $50 million received from delivering the note funds the loan that allows XYZ Co. to reverse in the note, and XYZ Co. has net interest earnings of $6,805.55 ($6,805.55 = $9,722.22 interest income less $2,916.67 interest expense). More generally, a very low specials rate presents an opportunity to speculate on an increase in the rate—by lending on an RP with the intention of failing at least initially—with limited downside exposure.

The supply schedules of securities available for lending also contracted as a result of a prospective increase in the operational costs of lending. After outright trading in Treasury securities reopened on Thursday, September 13, market participants became aware of the substantial volume of fails precipitated by the events of September 11. They realized that if they lent Treasury securities on RPs, they might not receive their securities back promptly at the end of the agreements. This would leave them exposed to the risk of counterparty failure and necessitate the allocation of scarce resources to monitor and control that risk. (The withdrawal of some institutional lenders from the specials markets after September 11 was widely noted. See, for example, *Dow Jones Newswire* [2001a, b], *Bloomberg* [2001b, in which Peter Fisher, Treasury's Under Secretary for Domestic Finance, remarked on the "reluctance by institutional investors to lend into a market that is suffering from extraordinarily high fails levels"], *Wall Street Journal* [2001b], *New York Times* [2001], and *Washington Post* [2001].)

Despite painstaking efforts by market participants to reconcile their transactions, and despite the expanded volume of securities borrowings from the System Open Market Account, settlement fails in the Treasury market did not abate quickly to pre-attack levels.

Some market observers also suggested that a nontrivial quantity of Treasury securities may have been acquired in a "flight to quality" after the attacks by investors who normally invested in other types of securities and who were not prepared to lend the Treasury securities that they acquired (*Bloomberg* 2001a; *Wall Street Journal* 2001c). Any such migration of ownership would have further reduced the supply of securities available for lending at any given specials rate.

Comment

The inverse relationship between settlement fails and the quantity of securities available for lending at a given specials rate suggests that there may be multiple equilibria in the specials markets. The surge in fails on and immediately after September 11 may have moved the market to a new equilibrium characterized by higher levels of fails, lower levels of securities lending at any given specials rates, and lower equilibrium specials rates. Market participants were not ignorant of the possibility that the specials markets may have arrived at a new equilibrium in late September. A broker was quoted as wondering whether "desks aren't lending because they're scared [trades] won't settle," or whether "trades won't settle because nobody is lending?" (*Dow Jones Newswire* 2001a). Under Secretary Fisher later described the relationship between fails and reluctance to lend as "self-compounding" (*New York Times* 2001).

Subsequent Federal Reserve and Treasury Actions

The persistence of settlement fails in late September led the New York Fed to continue to relax restrictions on the SOMA securities lending program. On September 18, it reduced the minimum fee for borrowing securities from 150 basis points to 100 basis points, and on September 27, it increased the amount of a security that it was prepared to lend to 75 percent of the amount owned. As shown in Chart 4, the higher lending limit led to a sharp increase in aggregate borrowings, from $6.4 billion on September 26 to $13.4 billion on September 27. Nevertheless, specials rates for the on-the-run five- and ten-year notes remained close to zero and, as noted above, the volume of fails remained unusually high, averaging $142 billion per day for the week ending October 3.

Reopening the Ten-Year Note

At 11 a.m. on Thursday, October 4, the Treasury took the unprecedented step of announcing that it would sell, at 1 p.m. that day, an additional $6 billion of the on-the-run ten-year note (the 5 percent note maturing August 15, 2011), thereby increasing the outstanding supply of the note to $18 billion. This was the first unscheduled snap auction of a coupon-bearing security since the regularization of note and bond auctions in the late 1970s and early 1980s.[18] The new securities were issued shortly after 9 a.m. the following morning.

The immediate reason for Treasury's unusual response to market conditions was a desire to reduce the volume of fails; Treasury attributed the decision to the "chronically high fails rate" (*Bloomberg* 2001a). More fundamentally, officials had become concerned that fails might begin to affect the performance and efficiency of the market. Under Secretary

Fisher stated that "we want to reduce the risk that these settlement problems turn into a much bigger problem for the Treasury market" and that "we wanted to prevent technical problems in the back office from causing wider problems in the pricing of government securities" (*New York Times* 2001; *Washington Post* 2001).

The additional ten-year notes, and the prospect that Treasury might also reopen the on-the-run five-year note,[19] appear to have had the intended effect. As shown in Chart 5, daily average fails fell from $142 billion during the week ending October 3 to $63 billion the next week and then to about $18 billion in each of the following two weeks. The specials rate for the ten-year note jumped to 90 basis points on October 5 and continued to rise to 173 basis points on October 12. The specials rate for the five-year note rose more modestly, but

At 11 a.m. on Thursday, October 4, the Treasury took the unprecedented step of announcing that it would sell, at 1 p.m. that day, an additional $6 billion of the on-the-run ten-year note.

reached 40 basis points by October 12. As shown in Chart 4, aggregate borrowings of securities from SOMA also declined after the reopening.

Comment

The impact of the reopening on the specials rate for the five-year note and on the general level of settlement fails is consistent with the idea that there may be multiple equilibria in the specials markets. In particular, the highly unusual and widely publicized Treasury action may have moved the market back toward an equilibrium characterized by a lower level of fails, a higher volume of securities lending at any given specials rate, and higher equilibrium specials rates. Market participants may have come to the conclusion that Treasury would take unprecedented steps to resolve the fails problem and that there was less reason to continue to hold back on lending securities out of concern that the securities might not be returned promptly.

Alleviating Chronic Fails

Settlement fails can become widespread and persistent when specials rates are near zero, as we discussed. Demand to borrow securities may substantially exceed lending supply at a specials rate just above zero, but the option to fail precludes an equilibrium rate below zero. Initiatives to alleviate chronic fails must therefore either increase an issue's lendable supply when specials rates approach zero or increase the cost of a fail and allow specials rates to go below zero.

Reopening Issues on an Unscheduled Basis

One way to alleviate chronic fails in an issue is to increase the issue's outstanding supply. Treasury adopted this approach when it reopened the ten-year note in October 2001. The reopening increased the lendable supply of the note, causing the specials rate on the issue to rise and making borrowing the note more attractive than failing.

Although the reopening seemed effective at resolving fails in the ten-year note (as well as in other issues), ad hoc reopenings have drawbacks. Above all, such reopenings do not necessarily meet Treasury's financing needs. As noted, Treasury indicated that the reopening of the ten-year note had "nothing to do with an increase of funding needs on our part." Financing the federal government is, of course, the purpose of debt issuance in the first place. Selling securities to enhance market efficiency could be seen as a matter of the tail wagging the dog.

Unscheduled reopenings also disrupt Treasury issuance patterns, both with respect to the reopening itself and with respect to any subsequent disruptions resulting from the additional funds. Since the early 1980s, Treasury has maintained a regular and predictable schedule of note and bond auctions (Fisher 2002b). Predictable offerings are believed to help minimize Treasury borrowing costs by reducing uncertainty among market participants.

Perhaps because of these drawbacks, Treasury has indicated that unscheduled reopenings are unlikely to become a prominent debt management tool. In early 2002, Under Secretary Fisher said, "never is a long time, so it would be imprudent of me to say that the Treasury will never again hold such an auction. But you should not count on it. . . . We want to rely on [market participants] to reconcile the forces of supply and demand" (Fisher 2002a).

Expanding Federal Reserve Securities Lending

Another way to alleviate chronic fails is to foster lending by large owners. Treasury securities owned by the Federal Reserve's System Open Market Account are in fact made available for overnight borrowing, as we discussed. However, the ability of SOMA lending to mitigate settlement problems is constrained by limits on the quantities of securities available for lending. As shown in Charts 6 and 7, such limits were binding on fifty trading days in 2001 for the on-the-run five-year note and on seventy-nine days for the ten-year note. (We classify the lending limit as binding when the amount borrowed is within 1 percent of the limit.)

The temporary relaxation of program rules on September 27 that made 75 percent of SOMA holdings available for lending allowed securities borrowing to expand.[20] Nonetheless, even the higher limit was sometimes binding and it was unable to move specials rates much above zero for some securities. SOMA lending is ultimately constrained by SOMA ownership. In September 2001, SOMA owned only $1 billion of the $12 billion on-the-run ten-year note.[21] The temporary increase of the SOMA lending limit from 45 percent to 75 percent thus allowed for additional lending of only $300 million of that issue.

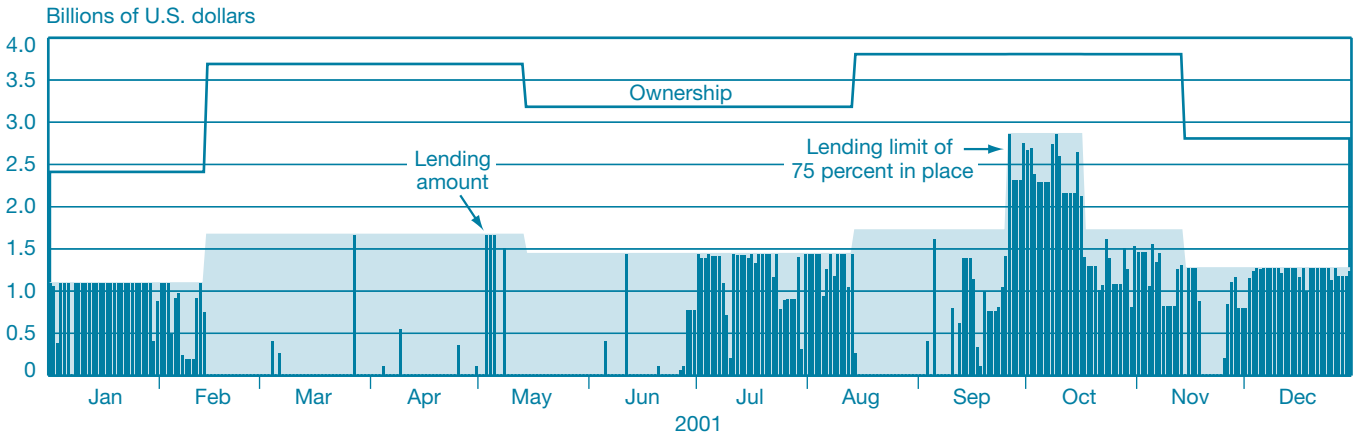
Introducing Treasury Securities Lending

As Treasury has the ability to create new securities, it could alleviate chronic fails by introducing a securities lending facility of its own. Such a facility was recommended by the Treasury Advisory Committee in the aftermath of September 11 (*Bloomberg* 2001c). The facility would enable market participants to borrow specific issues on a temporary basis, thereby expanding the supply of a security when it is in particularly high demand without permanently increasing its supply. As with the SOMA program, Treasury could require that any borrowed securities be collateralized with other Treasury securities.

In contrast to the SOMA program, a Treasury facility need not set limits on the quantity of securities it lends out. Treasury might instead want to specify the fee at which securities can be borrowed and then fill all applications at that fee. An appropriate fee might be one that was just below the general collateral rate, such as the general collateral rate less 10 basis points. With borrowing collateralized by other Treasury securities, such a fee would be equivalent to Treasury lending securities (and borrowing money) at a specials rate of 10 basis points and then relending the money (and borrowing other securities) at the general collateral rate. The fee would introduce a near infinite elasticity of supply of securities at a specials rate just above zero, but make borrowing securities from Treasury unattractive as long as specials rates were greater than 10 basis points.

Chart 6

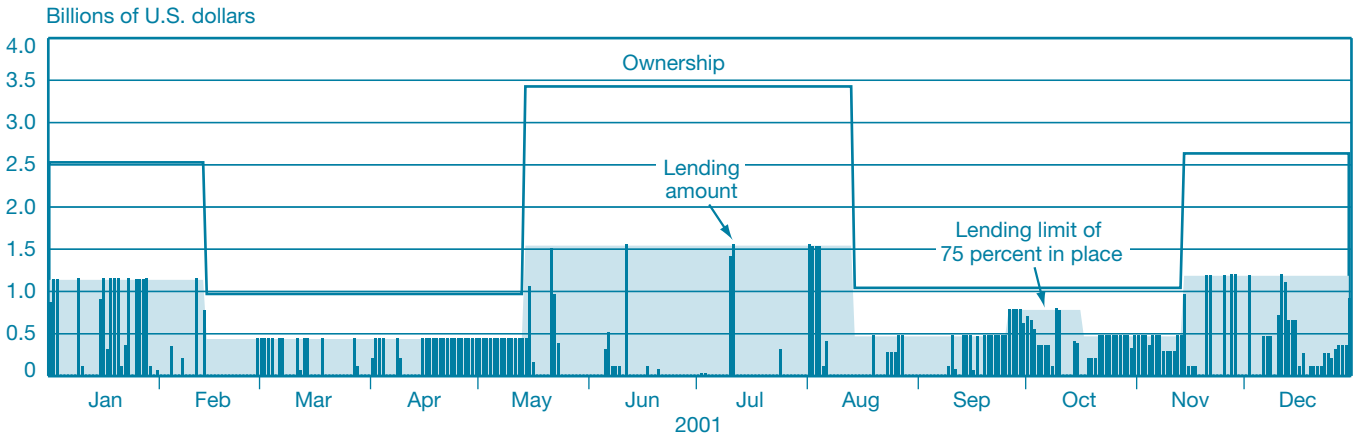
SOMA Lending versus Lending Limits for the Five-Year Treasury Note



Source: Federal Reserve Bank of New York.

Notes: The chart plots daily lending of the on-the-run five-year U.S. Treasury note by the System Open Market Account (SOMA) of the Federal Reserve against SOMA securities lending limits and ownership from January 2 to December 31, 2001. The shading indicates the lending limit.

Chart 7
SOMA Lending versus Lending Limits for the Ten-Year Treasury Note



Source: Federal Reserve Bank of New York.

Notes: The chart plots daily lending of the on-the-run ten-year U.S. Treasury note by the System Open Market Account (SOMA) of the Federal Reserve against SOMA securities lending limits and ownership from January 2 to December 31, 2001. The shading indicates the lending limit.

A fee marginally less than the general collateral rate would preserve the incentive to borrow an issue rather than fail without affecting the normal functioning of the specials market. Market participants would have an incentive to borrow a security from Treasury only when the specials rate dropped below 10 basis points, so issue specialness would likely be only minimally affected.[22] To the extent that issue specialness did not decrease, Treasury borrowing costs would not be adversely affected.[23] In fact, the mitigation of settlement problems could enhance market liquidity and efficiency, leading to lower Treasury borrowing costs.

In addition to the SOMA program, there is precedent for a Treasury securities lending facility from other debt issuers, as described below. Table 2 compares key features of the various lending facilities.

United Kingdom Debt Management Office RP Facilities

The United Kingdom Debt Management Office has had a "special" RP facility available for several years. The facility allows the office, at its sole discretion, to create and repo out gilts to "dissuade manipulation" of specific issues and to "address instances of market disruption or dislocation when a particular stock is temporarily in extremely short supply" (United Kingdom Debt Management Office 1999). Such operations would likely allow market makers to borrow specific issues overnight against lending money at a zero percent rate, while simultaneously requiring them to borrow funds against general collateral at the Bank of England's RP rate (to offset the cash flow implications of the special RP).[24] The special RP facility has never been utilized, although it was made available on one occasion (Bank of England 1998).

The Debt Management Office introduced a nondiscretionary standing RP facility in June 2000 that provides for the creation and lending of any issue on demand and that largely supersedes the special RP facility (United Kingdom Debt Management Office 2000). The facility is available to market makers willing to lend money at a rate equal to 10 percent of the Bank of England's RP rate against borrowing a specific issue, and requires that they simultaneously borrow money (against general collateral) at the bank's RP rate. A market maker thus effectively pays a fee equal to 90 percent of the bank's RP rate to borrow a specific issue, while providing generic securities as collateral. The facility was first used on December 29, 2000, and has been used several times since (United Kingdom Debt Management Office 2001, p. 35).

Fannie Mae RP Facility

In the United States, the Federal National Mortgage Association (Fannie Mae), a government-sponsored enterprise, introduced an RP lending facility in October 2001. The facility is intended to enhance the liquidity and trading activity of Fannie Mae's securities in the cash and collateral markets "by reducing the possibility of fails and providing collateral when there is limited lendable supply" (Fannie Mae 2001, p. 1).

Table 2
Comparison of Securities Lending Facilities

Category	System Open Market Account Securities Lending Program	United Kingdom Debt Management Office Standing RP Facility	Fannie Mae RP Facility	Treasury Securities Lending Facility
Introduced	April 1999[a]	June 2000	October 2001	Proposed
Offering process	Daily auction	Fixed price offering	Daily auction	Fixed price offering
Fee/rate	Minimum fee of 100 bp	Fixed fee of 90 percent of the Bank of England's RP rate[b]	Maximum rate of GC-100 bp[c]	Fixed fee such as GC-10 bp
Collateral	Other securities	Other securities	Cash	Other securities
Lending limit	45 percent of holdings	None	60 percent of holdings	None
Term	Overnight	Overnight	Overnight	Overnight

Sources: Federal National Mortgage Association (Fannie Mae); Federal Reserve Bank of New York; United Kingdom Debt Management Office.

Note: Information for the three existing facilities is as of December 31, 2001.

[a]The program was implemented in its current form at this time but was first established in 1969.
[b]This is the implicit fee; the facility provides for securities lending at 10 percent of the Bank of England's repurchase agreement (RP) rate while requiring market makers to put up general collateral (GC) at the bank's RP rate against the issues being lent.
[c]Such a requirement is roughly equivalent to a minimum 100-basis-point (bp) fee in the other programs.

The program works as follows. Fannie Mae issues to its own account an additional 25 percent of each new "benchmark" issue brought to market. (Fannie Mae's benchmark securities are issued in large sizes, on a regular basis, and across a range of maturities.) Fannie Mae then lends out, on a daily basis, up to 60 percent of the amount of any security held in its own account. For example, if Fannie Mae brings a $5 billion issue to market, it retains an additional $1.25 billion of the issue, and makes up to $750 million of the issue available for lending on any single day. The RPs bring in funds for Fannie Mae, which it presumably relends in the general collateral market or uses in lieu of borrowings from other sources.

The amount lent out and corresponding rate are determined through a multiple price auction, with the highest acceptable bid rate based on the level of market rates. In December 2001, for example, the highest acceptable bid rate was set at the general collateral rate less 1.00 percent. (If the general collateral rate were 1.75 percent, such a requirement called for dealers to bid 0.75 percent or less.) Fannie Mae's RP facility was first used on November 26, 2001, and has been used numerous times since (Fannie Mae 2001, p. 2).

Instituting a Penalty Fee for Fails

Chronic fails can also be alleviated by increasing the cost of failing with a penalty fee. The penalty would give sellers an economic incentive to borrow securities, even at specials rates below zero, in order to make delivery. Specials rates below zero would also draw lenders into the market that might otherwise be unwilling to lend (because of inadequate compensation or because of concern that their securities might not be returned promptly) and would curtail borrowing demand arising out of short sales. Provision for market clearing specials rates below zero would also bring the Treasury market into line with U.S. equity markets, where stock loan rates can, and do, go below zero (D'Avolio forthcoming), and would be consistent with Under Secretary Fisher's intention "to rely on . . . the forces of supply and demand."

The penalty fee could possibly be instituted through good-practice recommendations of the Bond Market Association. Failed deliveries would be rescheduled for the next business day at a slightly lower invoice price (under certain circumstances), as opposed to an unchanged invoice price (the current practice). The operational burden of changing an invoice price following a delay in settlement would undoubtedly be substantial, but the burden could be limited by restricting the penalty to aged fails, such as those outstanding a week or more. Restricting the penalty to aged fails would also avoid penalizing fails arising from operational problems or miscommunication, which tend to get resolved quickly in any event.

The penalty fee could be set high enough so that failing is rarely an attractive option, but not so high as to unduly punish fails attributable to operational problems and/or lead to

protracted conflict between buyers and sellers over the specific causes of fails. For example, suppose the penalty fee were set at a rate of 5 percent per annum minus the general collateral rate, with a minimum of zero. The fee would then be implemented only at general collateral rates of less than 5 percent and only to bring the total cost of failing to 5 percent. If the general collateral rate were 3 percent, for example, the penalty fee would be 2 percent. Failure to deliver in such circumstances would result in settlement being rescheduled for the next business day, with the invoice price reduced at the annualized penalty rate of 2 percent. The failing party would incur an explicit cost of failing of 2 percent per annum in addition to the implicit cost of 3 percent per annum due to the delayed receipt of funds. Failing would not be an attractive option as long as the specialness spread on a security was less than 5 percent.

Conclusion

The attacks of September 11 destroyed the offices of several inter-dealer brokers in the Treasury securities market, impaired telecommunication services in and through Lower Manhattan, and precipitated an extraordinary volume of settlement fails in Treasury securities. The Federal Reserve responded by relaxing restrictions on its securities lending program, thereby making Treasury collateral more readily available. As connectivity was gradually restored, market participants began reconciling incomplete and erroneous trade information, and the volume of fails started to decline.

Despite painstaking efforts by market participants to reconcile their transactions, however, and despite an expanded volume of securities borrowings from SOMA, settlement fails did not abate quickly to pre-attack levels. The persistence of high levels of fails can be most immediately attributed to specials rates at or close to zero in the days and weeks following September 11. Specials rates were driven down by a contraction in lending associated with a lower fed funds rate and as a result of a prospective increase in the operational costs of lending. With specials rates so low, a fail was not an unattractive alternative to borrowing securities to make delivery.

The persistence of settlement fails ultimately led the Treasury to announce an unprecedented snap reopening of the on-the-run ten-year note. The reopening increased the note's specials rate and aggregate fails quickly declined. However, such reopenings are unlikely to become a recurrent debt management tool. Alternative, less disruptive, ways of alleviating chronic fails include the introduction of a securities lending facility run by Treasury and the institution of a penalty fee for fails.

Endnotes

1. Bank for International Settlements (2001) offers an excellent glossary of many of the terms that appear in this article.

2. The criteria are described at <http://www.newyorkfed.org/bankinfo/regrept/primary.html>. The primary dealers are listed at <http://www.newyorkfed.org/pihome/news/opnmktops/2002/an020401.html>.

3. There is some double counting in the figures, especially for brokered trading, as trades between primary dealers are reported by both parties. Trading volume statistics are posted at <http://www.newyorkfed.org/pihome/statistics/>.

4. JPMorgan Chase Bank resulted from the merger of the Chase Manhattan Bank and Morgan Guaranty Trust Company of New York on November 10, 2001. Clearing bank services were provided by the Chase Manhattan Bank prior to the merger. For expositional clarity, we refer to JPMC even when speaking of events prior to the merger.

5. The Federal Reserve limits a single Fedwire transfer to no more than $50 million (Board of Governors of the Federal Reserve System 2002, Section 10.2). The market practice of delivering larger sales in multiple separate blocks of $50 million (plus a tail piece, if required) is based on the delivery guidelines of the Bond Market Association (Public Securities Association 1993, Chapter 7, Section A.4.b).

6. Membership requirements are described in the GSCC Overview and the GSCC Rulebook, both of which are posted at <http://www.gscc.com/important_notices_frame.html> under Other Important Documents. Membership is posted at <http://www.gscc.com/participants_frame.html>.

7. This net obligation also includes forward settling trades executed on preceding days and scheduled to settle on the following business day and (as described in the next section) closing legs of repurchase agreements terminating on the following business day. It does not include trades executed earlier in the same day for cash settlement. Cash settlement trades are settled bilaterally during the business day, even if both the buyer and the seller are GSCC members, because GSCC does not assess net receive and deliver obligations until after the close of business.

8. In the collateral markets, a security is on the run if it is the most recently *issued* security in a particular series (as opposed to the most recently auctioned security), because securities lending cannot take place until a security has been issued.

9. In "Announcement of Revisions to the SOMA Securities Lending Program," Federal Reserve Bank of New York, February 12, 1999, posted at <http://www.newyorkfed.org/pihome/news/announce/1999>. Details on the securities lending program as it existed prior to September 11 also appear in announcements dated April 20, August 26, and September 7, 1999.

10. However, the Master Repurchase Agreement suggested by the Bond Market Association (posted at <http://www.bondmarkets.com/market/agreements.shtml>) provides, inter alia, that the lender of money on an RP may, at its option, declare an event of default following a failure by the borrower of money to deliver securities on the starting leg of the RP, and that the borrower of money may, at its option, declare an event of default following a failure by the lender of money to deliver securities on the closing leg of the RP (see paragraph 11 of the agreement). Outright transactions do not usually have comparable express default provisions, but a buyer may specify at the time a transaction is negotiated that it needs "guaranteed delivery."

11. GSCC marks fails to market every day and thereby eliminates the replacement cost risk of a continuing fail. The invoice price specified in an instruction to deliver a security to GSCC is revised to the contemporaneous current market value of the security each day that a member continues to fail to deliver the security. The change in invoice price is offset with a cash payment called a "fail mark." If the CMV increases, the member has to make a payment equal to the product of the increase in the CMV and the quantity of securities to be delivered. (The member recovers this payment when it delivers the security at the revised, higher CMV.) Conversely, if the CMV declines, the member will receive a cash payment. At the same time, GSCC marks to market its own fail to the member that was issued the matched and bound redelivery instruction.

12. See GSCC Rulebook, Rule 12, Section 4 (January 10, 2002). However, GSCC does have the right, at its sole discretion, to renet and pair off fails. See GSCC Rulebook, Rule 11, Section 8 (January 10, 2002). GSCC exercises its right of pair-off on opening-leg fails of blind-brokered RPs (GSCC Important Notice GSCC054.96, dated July 9, 1996, p. 4). In April 2002, GSCC introduced an automated pair-off process for members that have receive and deliver obligations for identical principal amounts of the same security (GSCC Important Notice GSCC018.02, dated March 1, 2002). GSCC important notices are posted at <http://www.gscc.com>.

13. However, some participants have, on rare occasions, reversed in securities at even negative specials rates when a settlement fail would entail an additional penalty. Such penalties can arise in failing to settle short futures positions and guaranteed delivery sales.

14. This process was not flawless. Not all of the dealers had reported all of their brokered trades to GSCC, so GSCC failed to capture one side of some brokered trades and failed to capture other brokered trades entirely. The omissions were ultimately corrected in the course of a prolonged and painstaking reconciliation process.

15. In some cases, a GSCC member made good delivery to GSCC on September 11 but GSCC nevertheless recorded a fail in the absence of knowledge of the delivery (GSCC Important Notice GSCC068.01).

16. Such pair-offs were part of the original design of the program for blind-brokered RPs. See GSCC Important Notice GSCC054.96, p. 4.

17. See the discussion of marking fails to market in endnote 11. In some cases, GSCC marked fails to market erroneously and created erroneous revised delivery instructions because it did not know that members had in fact made good delivery on September 11 (GSCC Important Notice GSCC087.01).

18. Prior to 2001, Treasury overtly tailored the sale of a coupon-bearing security to market conditions on only one occasion—in November 1992, when it reopened the 6 3/8 percent ten-year note maturing August 15, 2002, "in order to alleviate an acute, protracted shortage" (Office of Financing press release, dated November 3, 1992). However, that action amounted to no more than a decision to reopen the ten-year note in a regularly scheduled auction in lieu of selling a new note. In contrast, the note sale on October 4, 2001, was a reopening in an unscheduled snap auction.

Between 1979 and 2001, Treasury conducted two snap auctions of cash management bills, on May 30, 1980, and April 6, 1993. Both auctions came promptly after Congressional action to extend or increase the statutory ceiling on the federal debt and were the result of Treasury's need for cash rather than an unusual demand for the securities by market participants (*Wall Street Journal* 1980, 1993).

19. Market participants had been aware of chronic settlement fails in the five-year note since late September (*Dow Jones Newswire* 2001a). At the time Treasury announced the reopening of the ten-year note, it stated that "we will observe how effective this is at trying to deal [with the fails problem]. We will then consider whether a similar extraordinary auction of five years is something we might do next week" (*Bloomberg* 2001b).

20. Although the 45 percent limit was reimposed after three weeks, the limit was subsequently raised to 65 percent on May 15, 2002. The May 2002 changes are posted at <http://www.newyorkfed.org/pihome/news/opnmktops/2002/an020424.html>.

21. One reason SOMA owned only $1 billion of the ten-year note was that the security had just been issued in August and was not scheduled to be reopened until November. After the November reopening, SOMA ownership jumped to $2.6 billion. (SOMA could not participate in the unscheduled October reopening because it did not have any contemporaneously maturing securities that it could roll over.)

22. It is possible, however, that the facility could have a more significant effect on specialness by reducing lenders' concerns regarding chronic fails, leading to an expansion in the supply schedule of securities available for lending and a reduction in the time period over which an issue trades near a specials rate of zero.

23. Duffie (1996) explains why issue specialness should be reflected in securities prices, and Jordan and Jordan (1997) present empirical evidence supporting the hypothesis. Such specialness can be expected to be incorporated into prices at auction, leading to lower Treasury borrowing costs (all else equal).

24. The Bank of England's RP rate is the official short-term rate at which the bank lends to the money market and accordingly implements monetary policy (Bank of England 1999).

References

Bank for International Settlements. 1992. "Delivery versus Payment in Securities Settlement Systems." Committee on Payment and Settlement Systems, September. Available at <http://www.bis.org/cpss/cpsspubl.htm>.

———. 2001. "A Glossary of Terms Used in Payments and Settlement Systems." Committee on Payment and Settlement Systems, July. Available at <http://www.bis.org/cpss/cpsspubl.htm>.

Bank of England. 1998. "Delivery of 9% Treasury Loan 2008 into the March Long Gilt Future Contract on LIFFE." News Release, February 16. Available at <http://www.bankofengland.co.uk/pr98017.htm>.

———. 1999. "The Transmission Mechanism of Monetary Policy." Monetary Policy Committee, April. Available at <http://www.bankofengland.co.uk/montrans.pdf>.

Bloomberg. 2001a. "U.S. to Sell $6 Billion in 10-Year Notes to Help Meet Demand." October 4.

———. 2001b. "U.S. Sells $6 Billion in 10-Year Notes to Help Overcome Shortage." October 4.

———. 2001c. "Treasury Advisory Committee Recommends Repo Facility Creation." November 1.

Board of Governors of the Federal Reserve System. 2002. "Operating Circular 7, Book-Entry Securities Account Maintenance and Transfer Services." Available at <http://www.frbservices.org/industry/pdf/Oc7.pdf>.

Bond Market Association. 2001. "Minutes of Emergency Meetings of the Bond Market Association, September 11-21, 2001." Available at <http://www.bondmarkets.com/market/9-11_ minutes.shtml>.

Boni, Leslie, and J. Chris Leach. 2001. "Depth Discovery in a Market with Expandable Limit Orders: An Investigation of the U.S. Treasury Market." Unpublished paper, University of New Mexico and University of Colorado, July.

Cornell, Bradford, and Alan Shapiro. 1989. "The Mispricing of U.S. Treasury Bonds: A Case Study." Review of Financial Studies 2, no. 3: 297-310.

D'Avolio, Gene. Forthcoming. "The Market for Borrowing Stock." Journal of Financial Economics.

DeGennaro, Ramon, and James Moser. 1990. "Failed Delivery and Daily Treasury Bill Returns." Journal of Financial Services Research 4, no. 3: 203-22.

Dow Jones Newswire. 2001a. "After Attack, Settlement Woes Still Clogging Repo Market." September 26.

———. 2001b. "Grim Repo Still Causing Headaches for Treasurys." September 28.

Duffie, Darrell. 1996. "Special Repo Rates." Journal of Finance 51, no. 2: 493-526.

Dupont, Dominique, and Brian Sack. 1999. "The Treasury Securities Market: Overview and Recent Developments." Federal Reserve Bulletin 85, no. 12: 785-806.

Fannie Mae. 2001. "Fannie Mae Initiates New Lending Facility to Enhance Repurchase Financing Market for Bullet Benchmark Securities." fundingnotes, November. Available at <http://www.fanniemae.com/markets/debt/pdf/fundingnotes_11_01.pdf>.

Federal Reserve Bank of New York. 2002. "Domestic Open Market Operations during 2001." Available at <http://www.newyorkfed.org/pihome/omo/omo2001.pdf>.

Fisher, Peter. 2002a. "Remarks before the Bond Market Association Legal and Compliance Conference." January 8. Available at <http://www.treas.gov/press/releases/po906.htm>.

———. 2002b. "Bloomberg News' Bonds Roundtable 2002." February 13. Available at <http://www.treas.gov/press/releases/po1010.htm>.

Fleming, Michael. 1997. "The Round-the-Clock Market for U.S. Treasury Securities." Federal Reserve Bank of New York Economic Policy Review 3, no. 2 (July): 9-32.

———. 2000. "Financial Market Implications of the Federal Debt Paydown." Brookings Papers on Economic Activity, no. 2: 221-51.

References (Continued)

———. 2002. "Are Larger Treasury Issues More Liquid? Evidence from Bill Reopenings." Journal of Money, Credit, and Banking 34, no. 3: 707-35.

Garbade, Kenneth. 1996. Fixed Income Analytics. Cambridge, Mass.: MIT Press.

Government Securities Clearing Corporation. 2002. GSCC – 2001 in Review. New York: Government Securities Clearing Corporation.

Jordan, Bradford, and Susan Jordan. 1997. "Special Repo Rates: An Empirical Analysis." Journal of Finance 52, no. 5: 2051-72.

Keane, Frank. 1996. "Repo Rate Patterns for New Treasury Notes." Federal Reserve Bank of New York Current Issues in Economics and Finance 2, no. 10 (September).

MacRae, Desmond. 2001. "Looking into the Abyss." Global Custodian, winter: 70-6.

Martin, A. E. 1985. "The Book-Entry System for Treasury Securities." Federal Reserve Bank of Atlanta Economic Review 70, no. 8: 15-6.

McAndrews, James J., and Simon M. Potter. 2002. "Liquidity Effects of the Events of September 11, 2001." Federal Reserve Bank of New York Economic Policy Review 8, no. 2 (November): 59-79.

New York Times. 2001. "U.S. Acts on Shortage of Treasuries." October 5, p. C1.

Public Securities Association. 1993. Government Securities Manual. New York: Public Securities Association.

Sender, Henny. 1986. "The Day the Computers Went Down." Institutional Investor, March: 203-7.

Stigum, Marcia. 1988. After the Trade: Dealer and Clearing Bank Operations in Money Market and Government Securities. Homewood, Ill.: Dow Jones-Irwin.

United Kingdom Debt Management Office. 1999. "Consultation Paper on DMO Policy Regarding 'Special' Gilt Repo Operations." September 17. Available at <http://www.dmo.gov.uk/publication/g2gil.htm>.

———. 2000. "Response to DMO Consultation Document on 'Special' Gilt Repo Operations." February. Available at <http://www.dmo.gov.uk/publication/g2gil.htm>.

———. 2001. "DMO Annual Review." September. Available at <http://www.dmo.gov.uk/publication/f2ann.htm>.

Wall Street Journal. 1980. "Treasury to Offer $16.2 Billion in Bills, Notes This Week." June 2, p. 31.

———. 1985. "A Computer Snafu Snarls the Handling of Treasury Issues." November 25, p. 58.

———. 1993. "Bond Prices Rise on Report that Treasury Study Will Urge Change in U.S. Borrowing Patterns." April 7, p. C19.

———. 2001a. "Treasury Market Is Faced with Incomplete Trades." October 3, p. C10.

———. 2001b. "Treasury Sale Averts a Crisis in 'Repo' Market." October 5, p. C1.

———. 2001c. "Treasurys Fall as Wall Street Urges the Treasury to Sell More 5-Year Notes to Ease 'Repo' Problems." October 10, p. C15.

———. 2001d. "Banks and Regulators Drew Together to Calm Markets after Attack." October 18, p. A1.

Washington Post. 2001. "In Surprise, Treasury Holds Auction of 10-Year Notes." October 5, p. E1.

Jason Bram, James Orr, and Carol Rapaport

Measuring the Effects of the September 11 Attack on New York City

- The total cost of the September 11 attack on the World Trade Center—comprising earnings losses, property damage, and the cleanup and restoration of the site—is estimated to be between $33 billion and $36 billion through June 2002.
- The earnings losses consist of $7.8 billion in deceased workers' prospective lifetime earnings and $3.6 billion to $6.4 billion in reduced wage and salary income in city industries affected by the attack.
- The cost of cleaning up the site, replacing the destroyed World Trade Center buildings, and repairing damaged buildings and infrastructure is expected to reach $21.6 billion.
- Although the loss of life and disruption of activity temporarily reduced New York City's productive capacity, the attack's effects on employment and consumer confidence had largely run their course by mid-2002.

The attack on the World Trade Center on September 11, 2001, traumatized New York City and the nation. Almost 3,000 lives were lost, and more than 30 million square feet of office space in Lower Manhattan was damaged or destroyed. The loss of workers, physical capital, and infrastructure reduced the productive potential of the city's economy and disrupted the lives of hundreds of thousands of people. Damage to the transportation and communications infrastructure depressed economic activity for a number of months, especially in Lower Manhattan.

This article evaluates the short-term economic consequences of the attack on Manhattan and the four other boroughs that make up New York City. We begin with the deepest loss—that of human lives. We then look at the effects of the attack on the inputs to the production process: labor and capital.

The attack led to an idling and underutilization of labor not only in the World Trade Center area, but also in other parts of the city. (Views of New York City and Lower Manhattan are provided in Appendix A.) Our analysis of labor focuses on aggregate city employment as well as on industry effects and factors that impact employee productivity, including health and confidence.[1] The analysis of capital covers the destruction of commercial space and infrastructure. We also discuss the effects of the attack on the markets for office space, home construction, and home sales. Finally, we examine how the attack affected the city's most economically vulnerable residents.

Jason Bram is an economist, James Orr a research officer, and Carol Rapaport an economist at the Federal Reserve Bank of New York.

The authors thank Simon Potter and Robert Rich for valuable input and insights, as well as two anonymous referees. They also acknowledge the excellent research assistance of Silvia Ellis, David Lagakos, and Alisdair McKay. The views expressed are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System.

This framework is an economic one, incorporating quality-of-life issues. To pursue our analysis, we have restricted ourselves to the labor and capital markets. In theory, it should be possible to evaluate output and income losses directly. In practice, however, such an evaluation is unworkable because official tabulations of gross New York City product do not exist and income figures are reported with a considerable lag.[2] Thus, output effects must be inferred from the behavior of the labor and capital markets. Whenever possible, we separate the effects of the attack from the effects of the business cycle (although we do not attempt to isolate the effects of the fall 2001 anthrax scares from the effects of the attack). Unless otherwise noted, the data presented here cover the period through June 2002, the end of the recovery process at the World Trade Center site.

We conclude that the attack disrupted New York City's economy in many ways. Although it is difficult to put a dollar value on lives lost, it is also inappropriate to omit loss of life from an estimate of the damage sustained. Our intention is to present as complete a picture of the attack's effects as possible. Accordingly, we estimate that the aggregate present value of lost lifetime earnings for these workers is about $7.8 billion. In addition, in the nine months following the attack, lost jobs and a reduction in the number of hours worked translated into an estimated shortfall in aggregate earnings of $3.6 billion to $6.4 billion. The cost of replacing the destroyed and damaged physical capital and infrastructure is estimated at $21.6 billion. Finally, the sum of these labor and capital losses yields an estimated total loss through June 2002 of between $33 billion and $36 billion.[3]

Loss of Life

The death of almost 3,000 people in the attack was a loss to New York City and to the nation. This number includes those who worked in the two World Trade Center towers, the firefighters and police personnel who responded to the attack, and the tourists and other visitors who were in the World Trade Center complex that morning. The method we use to value loss of life is based on the concept of "lifetime-earnings loss." This method estimates individual economic losses by adding up a worker's pretax annual income from the year of death to the year that he or she had expected to retire.[4] For those who died in the attack, the estimated earnings loss is calculated by multiplying the average expected level of annual earnings by the average number of years left to work before retirement.[5]

We estimate workers who died in the attack earned, on average, $127,000 a year. This estimate is based on the average income in 2000 for all workers in Manhattan and all workers in the finance and insurance sectors in Manhattan. The average annual income for workers in the finance and insurance sectors—where about half of the deceased workers had been employed—is estimated to be $197,275 in 2002. The average annual income of all workers in Manhattan, excluding the two sectors, is estimated at $57,000.[6] We use the average age of the workers killed in the attack, forty, and assume that they had twenty-two more years left to work until retirement. The average income of these workers is assumed to grow at the rate of inflation, which is assumed to equal the average discount rate. Under these assumptions, the current value of the aggregate earnings loss reaches about $7.8 billion, or an average of $2.8 million per worker.[7]

Although private insurance is expected to cover a portion of these losses, it is not likely that all of the workers had taken out private life-insurance policies. The earnings losses sustained by the workers' families will be partially covered by various charitable funds. In addition, the families of all World Trade Center attack victims will be eligible to receive compensation under the federal Victim Compensation Fund.[8] Although these various payments will partially offset losses to families and individuals, they do not reduce the overall cost of the attack because those payments represent costs to other parties, such as the government and insurance companies.

Employment Disruptions

In addition to the loss of lives, the attack on the World Trade Center had a dramatic disruptive effect on employment in New York City. The number of private-sector workers started to decline at the beginning of 2001 because of national and local business cycles. The level of employment bottomed out in March 2002 and edged up during the second quarter of the year (Chart 1). From the peak in employment in December 2000 to the trough in March 2002, the number of people working in New York City's private sector fell by 147,000, or 4.6 percent. (By comparison, the number of private-sector jobs lost during the 1989-92 recession was 344,000, or 11.4 percent.) In this section, we estimate the number of jobs lost because of the attack separately from those jobs lost because of the business cycle.

More than one-third of the net job losses in the recent downturn—specifically, 55,000 of the 147,000—occurred between January and September 2001. However, the sharpest

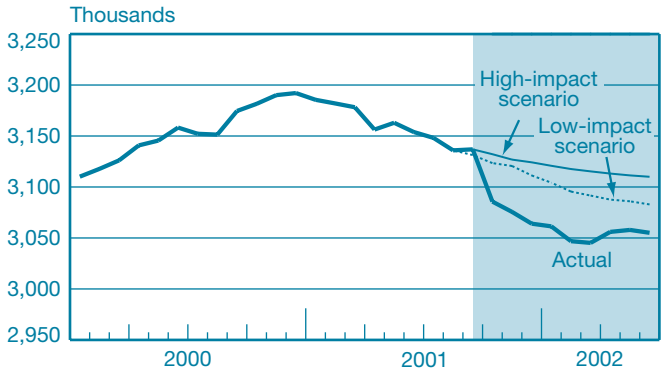
drop was in October 2001: a record 51,000 private-sector jobs were lost in that month alone. The remaining 41,000 job losses of the peak-to-trough decline occurred between October 2001 and March 2002. However, in the following months of April, May, and June, the number of private-sector jobs rose by a total of 10,000, or 0.4 percent.

To gauge how much of the fall in the number of jobs can be attributed to the attack, we use a standard dynamic forecasting model to estimate what the path of New York's employment would have been in the absence of an attack (Appendix B). The difference between the actual path of employment and this estimated path can be interpreted as the marginal effect of the attack on employment in the city at monthly intervals. Using this technique and two alternative sets of assumptions (high-impact scenario and low-impact scenario), we estimate that in October 2001, the number of private-sector jobs in the city was about 38,000 to 46,000 lower than it would have been otherwise. In February, this range moved to as high as 49,000 to 71,000, then eased to between 28,000 and 55,000 by June 2002 (Chart 2).

Data on weekly initial claims for unemployment insurance seem to confirm the pattern seen in payroll employment: the attack's effects on employment were substantial in October and November of 2001, but had largely run their course by early 2002 (Chart 3). Prior to September 11, weekly claims in New York City had been fluctuating in the 7,000 to 9,000 range—or about 1,000 to 3,000 higher than a year earlier, reflecting a general weakening in the economy. The weekly volume of claims more than doubled in the second half of September, and was running 10,000 to 12,000 higher than a year earlier, but then retreated steadily for four months, returning to approximate pre-attack levels by late February 2002. Aside from a brief spike in late March and early April—largely attributable to filings for extended benefits—the number of jobless claims was relatively steady throughout the first half of 2002.

These employment disruptions varied across the city's boroughs and neighborhoods, and across industries (Box 1). The most pronounced impact was concentrated in the blocks surrounding the World Trade Center, where numerous businesses, offices, and retail shops were either destroyed or

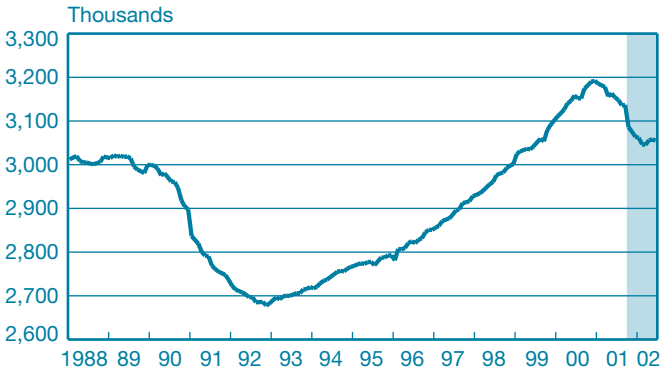
Chart 2
Path of New York City Private-Sector Employment



Sources: U.S. Department of Labor, Bureau of Labor Statistics; Federal Reserve Bank of New York.

Note: See Appendix B for methodology and a full explanation.

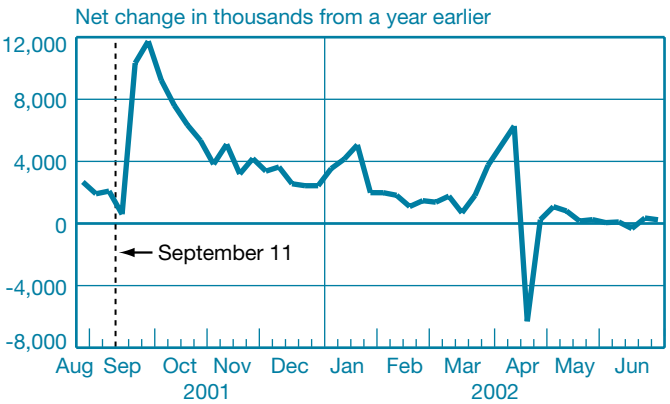
Chart 1
Private-Sector Employment in New York City



Source: New York State Department of Labor.

Note: The shading indicates the post-September 11 period.

Chart 3
New York City Initial Jobless Claims



Source: New York State Department of Labor.

Box 1

Employment Disruptions by Industry

The dynamic forecasting model suggests that most of the attack's net impact on employment levels occurred in October 2001. Here, we take a closer look at what appear to be the most directly affected industries: financial services, restaurants, hotels, and air transportation. Together, these industries accounted for 42,000 of October's 51,000 drop in private-sector employment. In subsequent months, although the estimated effect on *overall* employment was relatively modest, some industries registered further losses while others rebounded (see chart below). To get a better understanding of the attack's effects over time, it is helpful to examine these industries and their performance. Because swings in employment after September 11 are far larger than any preexisting trends within these industries, we assume that changes in employment after that date are mainly attributable to the attack.

The **financial services industry** appears to have been the most directly affected sector by far. In New York City, the number of jobs in the securities industry fell by 12,000, or 7 percent, in October 2001, and by an additional 6,000 from October 2001 to June 2002. In addition, the banking industry saw a net job loss of 8,000, or 8 percent, in October and lost another 1,000 jobs through June 2002. Net job losses in these key financial industries totaled 20,000 in October and another 7,000 through June 2002. Because some of the loss reflected a relocation of operations to nearby suburbs—mostly northern New Jersey—this figure overstates the net impact on the metropolitan area overall (see chart at right).

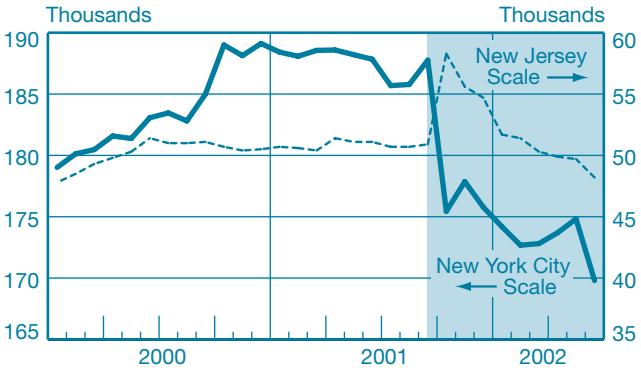
The **restaurant industry** also sustained steep job losses immediately following the attack. For the city overall, the number of jobs at bars and restaurants—which was imperceptibly affected at the national level—fell by an estimated 9,000 (6 percent) in October, but rebounded fully by December and held steady up to June 2002. However, these are net changes and do not capture the geographical distribution of employment in this industry. Thus, it is not clear if restaurant employment in the areas closest to the World Trade Center—the Financial District, Tribeca, and Chinatown—has fully rebounded to pre-attack levels.

The **hotel industry** lost an estimated 6,000 jobs, or 15 percent, citywide between September 2001 and March 2002. This reflected the drop-off in tourism, although 5,000 of those jobs were lost in October alone. In April 2002, the number of hotel jobs rose markedly by an estimated 4,000 and held steady in May at about 5 percent below pre-attack levels. Nationally, hotel industry employment has fallen by a more modest 4 percent since September 2001, but has yet to show any sign of bottoming out.

The steep decline in the number of people traveling also led to job losses in areas away from the World Trade Center site—in particular, at John F. Kennedy International Airport and LaGuardia Airport, both in the borough of Queens. The number of jobs in the city's **air transportation industry** fell by about 11,000, or 20 percent. Almost all of this decline occurred in October and November 2001, and there has been no sign of a rebound. Nationally, the number of jobs in this industry fell by 10 percent, with losses spread over the fourth quarter of 2001.

Although other industries, such as business services, apparel manufacturing, printing, and publishing, were also presumably affected, largely because of their strong concentration in Lower Manhattan, there is no indication of any significant shift in employment trends following September 11. However, it should be noted that many business owners and workers who did not lose their jobs evidently suffered income losses because of the disruptions in the weeks and months immediately following the attack. This is of particular concern in the restaurant and apparel industries, where workers' pay depends on business volume.

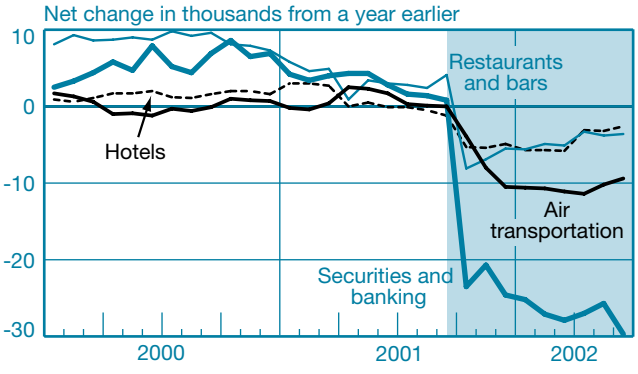
Financial Services Jobs in New York City and New Jersey
Seasonally Adjusted Level



Sources: New York State Department of Labor; New Jersey Department of Labor; Federal Reserve Bank of New York.

Note: The shading indicates the post-September 11 period.

Employment in Selected New York City Industries



Source: New York State Department of Labor.

Note: The shading indicates the post-September 11 period.

badly damaged. Substantial employment effects were also felt in the whole of Lower Manhattan (south of Canal Street [Appendix A]), where transportation access was curtailed and the volume of customer traffic fell precipitously. However, because of the drop-off in tourism—as well as possible multiplier effects from the loss of finance jobs—businesses throughout the city suffered because of the attack. For example, John F. Kennedy International Airport and LaGuardia Airport (both in the borough of Queens) saw, as did related businesses, a sharp decline in employment in the fourth quarter of 2001.

It is less clear whether the job losses were across all income levels. One might hypothesize that low-skilled, low-paid workers were more at risk of losing their jobs; labor economists generally maintain that the workers with the least job-specific skills are the first to be laid off in times of economic stress. Indeed, many of the workers in the hard-hit restaurant and retail sectors are relatively low-paid.[9] To test the hypothesis that the city's low-wage workforce faced a higher incidence of attack-related job loss than high-wage workers, we compare three industries where most employees are relatively well paid with three other industries where most employees are relatively poorly paid.[10] Both the high- and low-wage industries experienced a range of employment declines. Employees in the (low-wage) hotel and (high-wage) brokerage industries were especially affected. However, those in the (low-wage) general merchandise store and (high-wage) legal industries maintained previous employment trends. This example, although limited, does not support the hypothesis that the September 11 attack caused disproportionate job losses in low-wage industries.

We estimate that the attack led to a shortfall in wage and salary earnings of $3.6 billion to $6.4 billion as of June 2002. This estimate mainly reflects attack-related job losses but also includes the reduction in the number of hours worked.

The attack also led to a reduction in the number of hours worked. A recent study of the effects of the attack on workers in Chinatown indicates substantial short-term disruptions in the restaurant and garment industries.[11] Restaurants faced particularly severe declines in business volume in the weeks following the attack. These declines appear to have affected the number of hours worked as well as the number of jobs available. The garment industry also reported substantial declines in the number of hours worked (see Asian American Federation of New York [2002]).

On the basis of this analysis, we estimate that the attack led to a shortfall in wage and salary earnings of $3.6 billion to $6.4 billion as of June 2002. This estimate mainly reflects attack-related job losses, but also includes the reduction in the number of hours worked (Box 2).

Furthermore, worker productivity may have been lowered by changes in personal habits, health, and confidence. Vlahov et al. (2002) report the results from phone interviews with 988 adult Manhattan residents living south of 110th Street five to eight weeks after the attack. About 30 percent of the sample reported an increased use of cigarettes, alcohol, and/or marijuana. The same residents who increased their use of cigarettes and/or alcohol were also found to be more likely to have post-traumatic stress disorder (PTSD) and major depression. In a related study, Galea et al. (2002) report that about 7 percent of the phone sample reported psychological symptoms consistent with current PTSD and almost 10 percent reported symptoms consistent with depression. These percentages are about twice baseline values.

In addition, the New York City Department of Health and the Centers for Disease Control performed a door-to-door survey of 414 individuals living in the Battery Park City residential complex (next to the World Trade Center site) and two other downtown areas most directly affected by the attack (Centers for Disease Control and Prevention 2002). As of October 2001, almost 40 percent of the sample showed PTSD symptoms. Moreover, about 50 percent were still experiencing symptoms consistent with smoke inhalation from the still-burning fires.

Surveys of consumer confidence can also help shed light on the attack's psychological effect on behavior. The widely cited Conference Board survey is only available by census region (that is, New Jersey, New York, and Pennsylvania combined), but since 1997, Siena College in Loudonville, New York, has conducted a parallel monthly survey of New York State residents in which consumer confidence is reported separately for the New York City metropolitan area. According to the

[Consumer] confidence fell fairly sharply in September 2001, recovered somewhat in October, and then rebounded to above pre-attack levels in November.

Siena College (2002) report, the pattern of consumer confidence suggests a very short-lived effect from the attack. Confidence fell fairly sharply in September 2001, recovered somewhat in October, and then rebounded to above pre-attack levels in November. It remained well above its September trough through mid-2002. Interestingly, although this roughly parallels the national trend, U.S. consumer confidence did not begin to recover until December 2001, a month later than it did in the New York City area.

Overall, the effects of the attack were quite uneven across industries and workers. The finance, restaurant, hotel, and air transportation industries in the city were directly affected by the attack. Moreover, there is some evidence that the decline in business volume in Lower Manhattan (following a decline in demand) also led to a reduction in the number of hours worked, largely in the restaurant and garment industries. More generally, while many of the workers in the affected industries were relatively low-paid, we found no indication that employees in the city's lower paying industries were at significantly greater risk of losing their jobs because of the attack than were workers in higher paying industries. We did find some evidence, however, that the productivity of workers living in Manhattan may have been lowered in the immediate aftermath of the attack because of health problems. Nevertheless,

Box 2
Earnings Disruptions

To estimate the marginal effect of the attack on wage and salary earnings, we must first come up with a reasonable assumption regarding the average earnings per worker associated with the net job shortfall. Because the industry profile of attack-related job losses evidently differs from the city's overall industry mix, it would be inappropriate to assume that the average earnings associated with these job losses match the citywide average.

Although our employment simulation is based on a macroeconomic model that ignores the industrial profile of job losses, we can make assumptions about the mix of jobs lost based on total job losses by industry in the first few months after the attack (that is, October through December 2001). As indicated in the table, the most persistent job losses were concentrated in the financial services, air transportation, and hotel industries. The table shows two alternative estimates of the average earnings per worker in 2002 associated with the job shortfall. The "high-impact" scenario assumes that all of the job losses were concentrated in the financial, air transportation, and hotel industries. The "low-impact" scenario assumes that 75 percent of the job losses occurred in these industries, another 10 percent occurred in restaurants, and the remaining 15 percent was evenly distributed across all other industries.

These figures, combined with the employment scenarios described earlier, imply that total wage and salary earnings would have been between $3.4 billion and $6.2 billion higher if not for the attack. In addition, disruptions to Chinatown's garment industry and Lower Manhattan's restaurant industry may have reduced income by an additional $200 million, bringing the total estimated loss to within a range of $3.6 billion to $6.4 billion.[a]

	Distribution of Job Shortfall (Percent)		Average Earnings in 2002 (Dollars)	
Industry	Low-Impact Scenario	High-Impact Scenario	Low-Impact Scenario	High-Impact Scenario
Finance	45	64	197,275	197,275
Air transportation	10	13	50,752	50,752
Hotels	20	23	38,986	38,986
Restaurants	10	0	20,244	—
All other industries	15	0	61,511	—
Weighted average	100	100	115,470	142,775

Source: Authors' calculations.

Note: The 2002 average earnings figures are based on the 2000 County Business Patterns data for Manhattan (except for air transportation, where earnings are for Queens) and are increased by 8 percent.

[a] In the first few months after the attack, workers in Chinatown's garment industry reportedly incurred a steep fall-off in hours and income that was not reflected in the employment statistics (see Asian American Federation of New York [2002]). Although income data by industry are not yet available, aggregate reported income was about $220 million per quarter for the garment industry and $540 million for the restaurant industry in 2000. Our estimated $200 million earnings shortfall assumes a 25 percent reduction in hours and earnings (of those still employed) in these two industries persisting for one quarter.

Siena College's tracking of consumer confidence in the metropolitan area strongly suggests a mitigation of these adverse psychological effects and a general improvement in attitudes in subsequent months.

Physical Capital Losses and Damage

The major components of New York City's public and private physical capital stock in Lower Manhattan that were destroyed or damaged in the World Trade Center attack were as follows: about 30 million square feet of commercial office space and more than 100 retail stores in the World Trade Center area, subway tunnels (Lines 1 and 9), the Port Authority Trans-Hudson (PATH) train station at the World Trade Center, the streets surrounding the attack site, and parts of the telecommunications and power infrastructure in Lower Manhattan, including a switching facility and substations. In all, the resulting loss to the city's productive capacity is similar to what can follow an earthquake or major natural disaster.[12]

Several economic and financial measures have been used to estimate the dollar value of the city's physical capital losses associated with the attack.[13] In this article, we cite publicly available repair and replacement cost estimates for the major buildings and infrastructure affected by the attack. These dollar values are nominal gross replacement and repair costs over a multiyear period and do not explicitly account for the depreciation of the assets or any potential offsets from government rebuilding programs or private-insurance proceeds.

We group the main components of the city's physical capital losses directly related to the attack into three categories: 1) the cost of the cleanup and restoration for rebuilding at the site, 2) the cost of replacing about 14 million square feet of office and retail space in the World Trade Center complex and its contents and repairing the damaged buildings in the areas adjacent to the World Trade Center,[14] and 3) the cost of repairing the damage to the New York City subway lines, the destroyed PATH terminal in the World Trade Center, destroyed or damaged Con Edison facilities and equipment, and damaged telecommunications lines and equipment in Lower Manhattan.[15]

At the end of June 2002, the cleanup and restoration of the World Trade Center site was deemed complete and the final costs are expected to be about $1.5 billion (see table). These costs cover debris removal, street repair, police and firefighters' overtime pay, and other forms of disaster assistance and relief. Most of these expenses are expected to be reimbursed by the Federal Emergency Management Agency (FEMA).[16]

The cost of replacing destroyed or damaged buildings in the World Trade Center complex and adjacent areas is estimated to be $11.2 billion. Of this, $6.7 billion will be for rebuilding the destroyed World Trade Center complex, although it is unlikely that the pre-attack design will be duplicated.[17] The remaining $4.5 billion is the estimated cost of repairing the damaged buildings. The cost of replacing the contents of the destroyed buildings, including the technology and fixtures, has been estimated to be $5.2 billion.[18]

A tracking of former occupants in the World Trade Center complex shows that tenants from about 65 percent of the destroyed space have leased new space within New York City, with the majority relocating to midtown offices. Tenants from about 17 percent of the destroyed space have moved to New Jersey. It is expected that about two-thirds of the damaged property in the World Trade Center area will be reoccupied. It is also expected that tenants from about 11 percent of the damaged space will relocate to offices in New Jersey.[19]

The losses to the public infrastructure in Lower Manhattan are concentrated in three key areas—the collapsed subway tunnel and other damage to the 1 and 9 subway lines, the destroyed World Trade Center PATH station, and the damage to and destruction of parts of the telecommunications and power infrastructure. The Metropolitan Transportation Authority (MTA) has estimated the cost of repairing the subway lines to be $850 million and the Port Authority has estimated that restoring basic PATH service will cost $550 million.[20] FEMA funds can be used to meet these costs, although private insurance taken out by both the MTA and the Port Authority is expected to cover a portion of them.

The estimated cost of repairing the communications and power infrastructure is $2.3 billion, much of which is expected to be covered by private insurance and FEMA funds. Improvements to the infrastructure in Lower Manhattan will likely be undertaken, and the final bill, including these improvements, may well be significantly larger. The estimated total replacement and repair cost for these parts of the city's infrastructure is $3.7 billion. Although private insurance and funds allocated through FEMA will substantially offset much of the cost of these rebuilding efforts to New York City residents and businesses, the productive potential of the city was significantly reduced by the attack and will remain below its pre-attack level until the rebuilding is largely completed.

Aggregating the cost estimates for each of these components shows the total physical losses sustained in the attack to be about $21.6 billion.[21] To put this amount in perspective, it is equivalent to about 9 percent of the total earnings in New York City in 2000, or an average of $2,650 per

city resident. As we have observed, private insurance is expected to cover a significant amount of these losses, and FEMA funds appear to be sufficient to cover a substantial share of the uninsured public infrastructure costs. Of course, this coverage mitigates the cost to New York City residents but not to the nation as a whole.

These estimated replacement costs of the physical losses are based on the assumption that the reconstruction of the World

Impact of the World Trade Center Attack on New York City as of June 2002

Impact	Estimated Magnitude	Notes
Labor market		
Loss of human life	Estimated 2,780 workers, $7.8 billion lifetime-earnings loss	Losses estimated as present discounted value of lifetime earnings; federal Victim Compensation Fund set up to help offset earnings losses and psychological impacts on families
Net job losses	38,000-46,000 in October 2001, rising to 49,000-71,000 by February 2002, diminishing to 28,000-55,000 by June 2002	Most of the employment losses related to the attack were in finance, airlines, hotels, and restaurants
Net earnings losses	$3.6 billion to $6.4 billion between September 2001 and June 2002	Based on estimates of net job losses and reduced hours
Attack-related productivity effects	Some increase in post-traumatic stress disorder and alcohol and drug use three months after attack	Difficult to quantify attack's impact on workers' mental and physical disabilities
Total labor loss	$11.4 billion-$14.2 billion	
Physical capital		
Cleanup and site restoration	$1.5 billion	Completed June 2002; expenses covered by the Federal Emergency Management Agency (FEMA)
Destroyed buildings in World Trade Center complex	Approximately 14 million square feet, $6.7 billion to rebuild	Book value of towers at $3.5 billion; complex privately insured
Damaged buildings in World Trade Center area	Approximately 15 million square feet, $4.5 billion	Inclusion of damage to Class B and C space raises estimate to 21 million square feet
Contents of buildings in World Trade Center complex	$5.2 billion	Significant offset from private insurance
Public infrastructure		
Subway	$850 million	Estimated repair cost; significant offset from private insurance and/or FEMA for repair to all three components of infrastructure
PATH	$550 million	
Utilities	$2.3 billion	
Total capital loss	$21.6 billion	
Total (labor, capital) loss	$33 billion-$36 billion	

Notes: The rounding of the total (labor and capital) loss figure acknowledges imprecision in the estimates. On the one hand, estimates of the labor loss may be understated, primarily for two reasons: the June 2002 cutoff for estimating earnings impacts and the possible earnings reductions due to a drop in the number of hours worked (in industries other than apparel and restaurants). In addition, attack-related declines in worker productivity (due, for example, to stress) may have affected employed workers and are not captured in our estimated earnings losses associated with declines in employment and hours. On the other hand, estimates of the labor loss may be overstated, because of the double counting of the earnings losses of some of the deceased workers and the assumption that the deceased workers would have worked in New York City until retirement. Furthermore, although this earnings-loss tally corresponds to New York City proper, these figures will overstate the net impact on the broader metropolitan area and the nation because many of the job "losses" reflect job relocations from the city to the suburbs—largely northern New Jersey. Because these are aggregate loss estimates, the issue of distributional impacts is not addressed.

Trade Center area will essentially duplicate what existed before the attack. However, as of June 2002, a final reconstruction plan has not yet been reached and the subject remains under discussion.

The Lower Manhattan Development Corporation (LMDC), a public corporation with both city- and state-appointed members, is helping to coordinate the redevelopment of the site. The corporation has been soliciting from various advisory boards ideas for the redesign of the site, including putting a memorial to the attack victims on the site, setting aside part of the World Trade Center area for residential units, and reconfiguring the transportation linkages between PATH and the New York City subway lines. The ultimate cost of replacing the lost capital stock depends on the final decisions regarding redevelopment of the site.

Impact on the Office Market

One of the most dramatic and surprising outcomes of the attack was on Manhattan's (and the metropolitan area's) office market. Demand for office space had been weakening and vacancy rates rising prior to the attack. After the attack, with an estimated 3 percent of Manhattan's office space destroyed and another 3 percent rendered temporarily unusable, it was widely expected that a severe shortage of space would push down vacancy rates and cause a sharp spike in rents. However, quite the opposite occurred: vacancy rates rose further and rents declined (Chart 4). This happened because of a number of factors: demand weakened more than was anticipated, firms had a good deal of extra space (in both Manhattan and adjacent areas) that they were able to sublet to displaced firms, and some Manhattan hotels were retrofitted to serve as temporary office space.

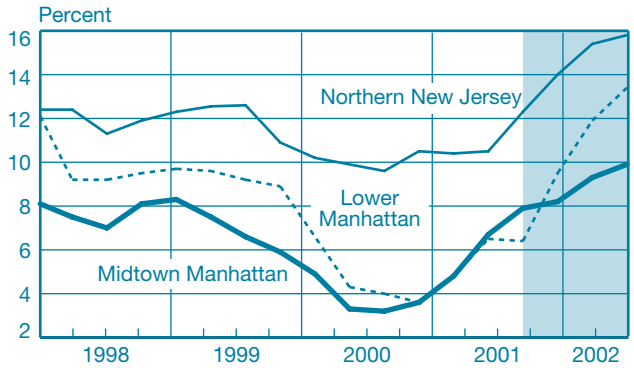
One of the most dramatic and surprising outcomes of the attack was on Manhattan's (and the metropolitan area's) office market. . . . Vacancy rates rose further and rents declined.

Impact on the Most Vulnerable

The preceding two sections focused on labor and capital losses. In this section, we look at the effects of the attack on the most economically vulnerable New York City residents.

Chart 5 shows the monthly aggregate number of public assistance caseloads and the Federal Reserve Bank of New York's index of coincident economic indicators since January 1999.[22] The bulk of public assistance is made through

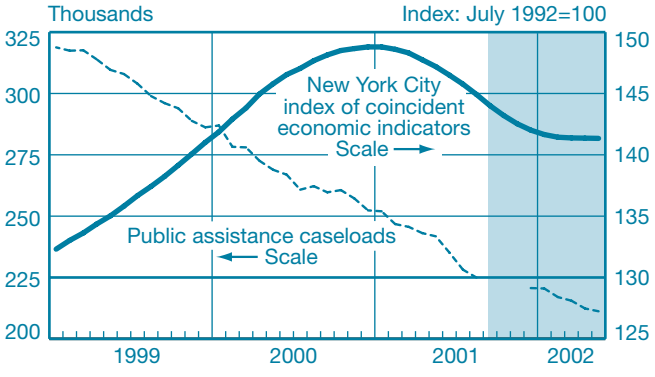
Chart 4
Office Vacancy Rates



Source: Cushman and Wakefield.

Note: The shading indicates the post-September 11 period.

Chart 5
Public Assistance Caseloads in New York City



Sources: Federal Reserve Bank of New York; New York City Human Resources Administration.

Note: The shading indicates the post-September 11 period.

Temporary Assistance to Needy Families, a federal and New York State block grant program. The remainder of public assistance includes the New York State programs Safety Net Assistance and Safety Net Non-Cash. The caseloads for these programs are evaluated together.[23]

Understanding the causes of a downward trend in welfare caseloads is notoriously difficult (Blank 2001). The decline in the number of caseloads observed in the city between January 1999 and August 2001 could have stemmed from economic expansion, the welfare reform incentives to reduce the number of caseloads, or both. Between January and August 2000, when the city economy was expanding, the number of public assistance caseloads fell 8.7 percent. Between January and August 2001, when the city's economy was contracting but the incentives for families to get off assistance were especially strong, the number of caseloads fell 10.7 percent. In short, the attack came at a time when the number of caseloads in New York City was falling rapidly, despite the slowing economy.[24] The post-September 11 data show that the downward trend in caseloads is stronger than the attack's effects.

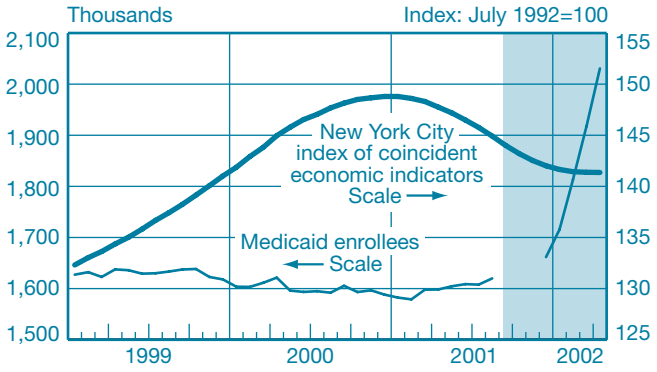
Chart 6 performs a similar exercise regarding the number of Medicaid caseloads. Medicaid—a federal government, New York State, and New York City matching entitlement program—provides medical assistance to certain low-income individuals and families with dependent children. Unlike public assistance, Medicaid enrollment displays some coincident sensitivity to the cycle. Between January and August 2000, when the city economy was expanding, enrollment fell by 12,000 to reach 1,592,000. Between January and August 2001, when the city economy was contracting, enrollment rose by 38,000. By December 2001, enrollment was up by 42,000, and by January 2002, it had reached 1,716,000.

The sharp increase in Medicaid enrollment after September 11 could stem from several factors. Those who were eligible for Medicaid but had not enrolled may have experienced worsening health from the attack and enrolled for the first time after September 11. In addition, those with incomes just above the Medicaid cutoff levels could have suffered attack-related income losses and become eligible.

However, the United Hospital Fund (2002) concludes that the increased enrollment is almost certainly the result of changes in the eligibility requirements for new enrollees. The attack disabled the Medicaid computer system and eligibility records, so New York City could not use the standard procedures to enroll patients. In response, the New York City Human Resources Administration and the New York State Department of Health developed a temporary assistance program, Disaster Relief Medicaid (DRM). DRM simplified the standard complex application process. Potential enrollees were asked only to fill out a one-page application stating that their income fell within certain guidelines. These individuals were then presumed to be eligible for DRM and received same- or next-day coverage.

Chart 6

Medicaid Enrollees in New York City



Sources: Federal Reserve Bank of New York; New York City Human Resources Administration.

Note: The shading indicates the post-September 11 period.

Summary of Losses

The loss of human life and the damage and destruction of commercial property and infrastructure that resulted from the September 11 attack significantly reduced the productive potential of the New York City economy. Moreover, the attack disrupted economic activity not only in the industries in the area of the World Trade Center, but also in a number of other industries throughout the city, further reducing employment.

In this article, we have assessed the impact of the attack on the city's economy by quantifying the effects on the inputs to the production process—labor and capital. We first considered the loss of human life. Although no single measure can capture the full impact of a premature death, the computation of the discounted value of a worker's expected future earnings is a conventionally used measure of an individual's economic loss. The attack claimed almost 3,000 lives and, using this discounted earnings measure, we estimate that it caused $7.8 billion in aggregate lost lifetime earnings for these workers and their families. This was as much a loss to the nation as to the city.

In addition, the attack caused significant declines in private-sector employment. Much of the job loss appears to have been concentrated in the finance, air transportation, hotel, and restaurant industries. Other adverse effects of the attack on the New York City labor market were also noted. In several industries, most notably restaurants and apparel, the hours worked by employees were significantly reduced. On the basis of these figures, the attack is estimated to have reduced city wage and salary income by a total of $3.6 billion to $6.4 billion. In addition, surveys found some increase in the incidence of PTSD and alcohol and drug use about three months after the attack, which likely resulted in time off from work and reduced productivity.

On the capital side, the attack caused an estimated $21.6 billion in physical capital and infrastructure losses. Adding this $21.6 billion in capital losses to the $11.4 billion to $14.2 billion in lost earnings yields a total loss of $33 billion to $36 billion. These losses include the costs of cleaning up the site, the replacement of the destroyed World Trade Center complex and its contents, the repair of the damaged buildings in the area, and the repair to the damaged public infrastructure. Although private insurance and FEMA funds are expected to cover a major portion of these costs, the loss of this capital is still a cost to the city's economy in terms of lost productive potential.

Recovery from the Attack

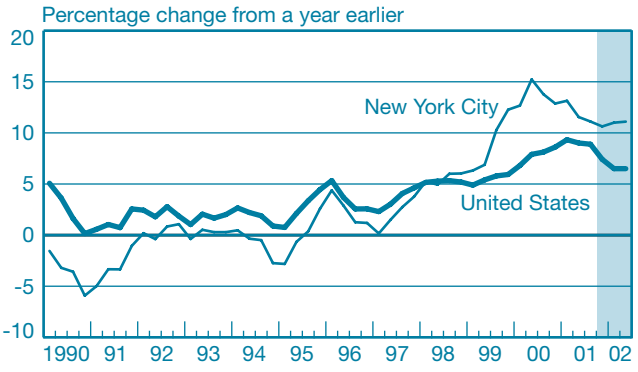
As we have observed, employment in the most clearly affected industries has been showing signs of a rebound since March 2002, despite little improvement at the national level and persistent weakness in the financial markets, which play a key role in driving the local economy. In terms of its distributional impact, the attack does not appear to have taken a strikingly disproportionate toll on low-skilled workers. Jobs in low-wage industries appear to have been adversely affected to the same degree as those in high-wage industries, and city welfare rolls show few signs of sudden growth in the months after the attack.

Moreover, while surveys have found some psychological harm to residents in the immediate area of the attack, consumer confidence in the metropolitan area had rebounded strongly as of mid-2002, suggesting that any widespread pessimism associated with the terrorist attack was short-lived in the New York City area, as it was nationwide. Another reflection of improved confidence can be seen in local housing markets. The market for Manhattan cooperative apartments and condominiums picked up noticeably in the second quarter of 2002, with the average selling prices rising an estimated 3 percent to 4 percent from a year earlier and the number of unit sales rising nearly 50 percent.[25] Similarly, selling prices of single-family homes in New York City's outer boroughs and nearby suburbs were an estimated 10 percent to 15 percent higher in the second quarter from a year earlier (Chart 7).[26]

Additional evidence of a recovery can be found by looking at the cleanup and restoration of the site, which was essentially completed months ahead of schedule and at a cost that appears to be substantially less than the amount of federal money allocated to the city for that effort. Furthermore, a number of programs have been established to support the relief and rebuilding efforts in Lower Manhattan. The Lower Manhattan Development Corporation, for example, was established in December 2001 to help coordinate the efforts to redesign and rebuild the World Trade Center area. In January 2002, a federal compensation program for the families of all victims of the attack, the first of its kind, was set up, and has since started making payments. Finally, the federal government has authorized grants, tax relief, subsidies, and other forms of assistance since the attack to aid in the rebuilding and redevelopment of Lower Manhattan.

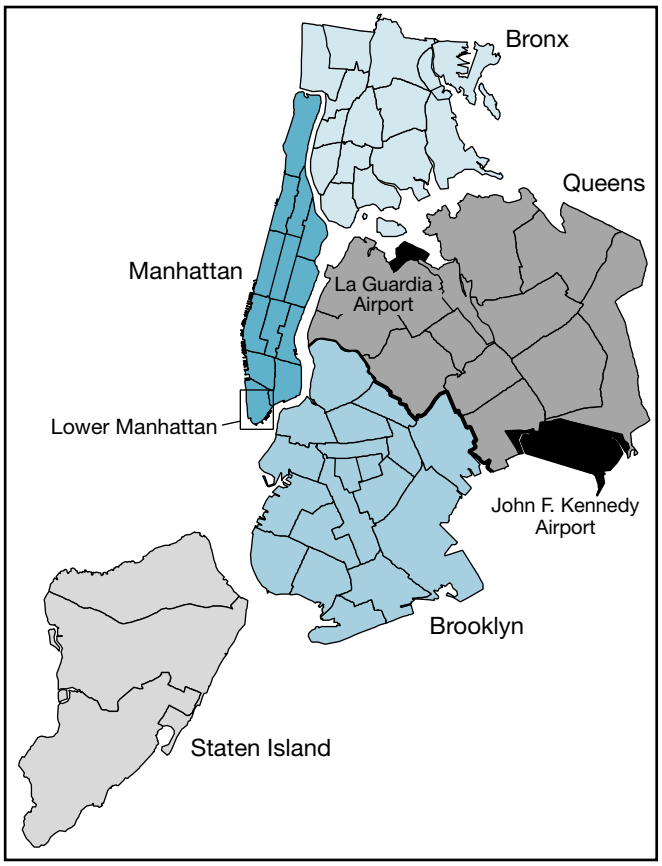
In conclusion, although New York City has clearly suffered a severe blow from the attack, the major disruptions appear to have been short-lived and conditions are in place to begin a recovery. At this point, the greatest challenge to the city comes from the economic fundamentals that have historically affected the local economy: the national business cycle and, in particular, developments in the financial markets.

Chart 7
Single-Family House Price Appreciation



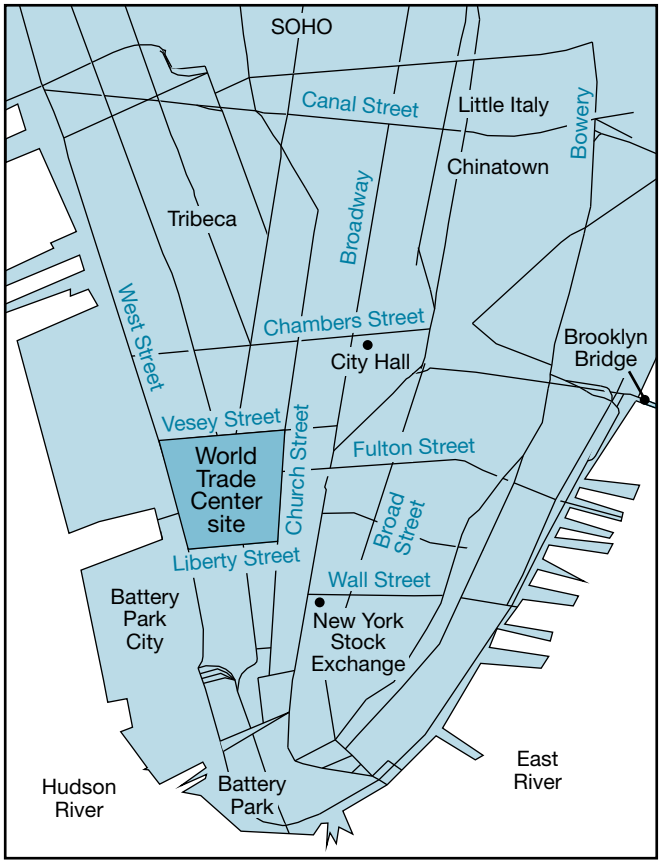
Source: Office of Federal Housing Enterprise Oversight.
Note: The shading indicates the post-September 11 period.

New York City Boroughs



Source: Federal Reserve Bank of New York.

Lower Manhattan



Source: Federal Reserve Bank of New York.

Appendix B: Methodology

Estimating the Effect of September 11 on the Path of Employment

To estimate the net impact of the terrorist attack on the subsequent path of employment, one must formulate a set of assumptions regarding the counterfactual—the path of employment had there been no attack. We do this by using an autoregressive forecasting model that estimates the relationship between employment growth in New York City and the rest of the nation. We then use this model to simulate the path of New York City employment after September 11 had there been no attack. However, there are various ways to specify this simulation, depending on a number of assumptions. To assess the robustness of the simulation (that is, to see how sensitive the outcome is to varying the assumptions), we run a number of simulations varying each of the following sets of specifications:

- *Number of lags used in the regression: three to eight.* The number of lags used in the estimation reflects the persistence of movements in employment. With relatively few lags, employment tends to snap back to its long-term trend relatively quickly, following any deviation. With more lags, employment reverts to trend more gradually. We run simulations using each of the above lag structures.

- *Post-September 11 U.S. data: actual, simulated.* Since it is generally preferable to use *actual* data whenever possible, we run one set of simulations using actual data for the United States (excluding New York City) after the attack. This implicitly assumes that employment outside New York City would not have behaved much differently if there had been no attack (that is, that the attack had a relatively small net effect on jobs outside New York City).

  If one assumes that the attack did have a significant impact on U.S. employment outside New York City, then using *actual* data after the attack would bias the results. Thus, we also perform a separate set of simulations in which U.S. data after September 11 are *predicted* based on pre-attack changes in employment for the private sector overall and for personnel-supply services. These estimates are then used in the original regression to predict New York City employment.

- *Last actual data point used: August, September.* Although the bulk of the effects of September 11 on employment showed up in the October 2001 data, it is possible that the September 2001 numbers were also slightly affected by the attack. Thus, we conduct one set of simulations using actual September data and another using actual data only through August.

As it turns out, the various combinations of assumptions yield results that do not vary dramatically. The weakest simulated employment path (low-impact scenario) is generated by using eight lags, with actual data for the United States, and using August as the last actual data point (for New York City). The strongest simulated employment path (high-impact scenario) is generated by using three lags, with predicted data for the United States over the simulation period, and using September as the last actual data point.

Equations:

e_t = private-sector employment in New York City,

E_t = private-sector employment in the United States (excluding New York City), and

P_t = U.S. personnel-supply employment.

Low-impact scenario (simulation begins after August 2001):

$$\hat{e}_t = \alpha + \sum_{i=1}^{8} \beta_i \hat{e}_{t-i} + \sum_{i=1}^{8} \gamma_i E_{t-i}.$$

High-impact scenario (simulation begins after September 2001):

$$\hat{e}_t = \alpha + \sum_{i=1}^{3} \beta_i \hat{e}_{t-i} + \sum_{i=1}^{3} \gamma_i \hat{E}_{t-i}$$

$$\hat{E}_t = \kappa + \sum_{i=1}^{3} \phi_i \hat{E}_{t-i} + \sum_{i=1}^{3} \eta_i \hat{P}_{t-i}.$$

Endnotes

1. Some of these factors affect residents who are unemployed.

2. Estimates of gross city product are reported by the New York City Office of the Comptroller. See <http://comptroller.nyc.gov>.

3. It should be noted that different concepts of losses are included in this sum, namely, replacement costs of capital, lifetime-earnings losses of the deceased workers, and the nine-month earnings losses of those idled because of the attack. Although there is evidently some double counting of losses in the latter two categories, we assume it is minimal and we make no adjustment for it.

4. The difficulties and pitfalls in putting a dollar value on human life are discussed in Dorman (1996). In addition to using a discounted earnings loss method, estimates of the economic value of a human life have also been based on observed wage premiums for job-related death risks faced by workers. A recent analysis using this methodology estimated the economic value of the life of an "average" worker to be between $1.5 million and $2.5 million in 1998 (Mrozek and Taylor 2002).

5. See the New York City Office of the Mayor for lists of the deceased (<http://home.nyc.gov>). The average age of those who died in the attack was 39.9 years.

6. These figures are based on data from the U.S. Department of Commerce and 2000 County Business Patterns. We obtain estimates for 2002 by incrementing those 2000 figures by 8 percent.

7. This method is similar to that used in the report by the New York City Office of the Comptroller (2001).

8. The federal Victim Compensation Fund was established by the federal government to compensate families of victims of the World Trade Center attack. A major component of the amount of compensation awarded to a family is the estimated lifetime-earnings losses of the victim adjusted for taxes, benefits, unemployment risk, and the victim's share of consumption. An additional sum is included in the compensation award for noneconomic losses. In calculating a victim's gross earnings losses, the fund assumes annual earnings increases of 3 percent from a combination of inflation and productivity growth, an annual increase related to experience (which rises at a decreasing rate), and a discount factor of 4.8 percent. Using these parameters, the fund estimates that a forty-year-old victim earning $127,000 would have lost $2.7 million. The ultimate compensation award is reduced by the amounts received from other sources of compensation, such as Social Security death benefits and life-insurance benefits.

9. Using a different methodology, the Fiscal Policy Institute (2002) concludes that the attack took a heavy toll on low-wage workers.

10. We used recent Bureau of Labor Statistics and Current Population Survey data to help identify two-digit industries where the average wages were toward the top or bottom of all New York City area industries. Within these groups, we selected three industries that represented a nontrivial fraction of city employment and displayed low wage variation across employees. We determined that hotels, food stores, and general merchandise stores are important low-wage industries in New York City, and engineering services, brokerage, and legal services are important high-wage industries. Although eating and drinking establishments is an important low-wage industry, wages varied across employees much more than they did in the selected industries.

11. Asian American Federation of New York (2002).

12. The most recent estimates of total insurance losses—including property, business interruption, aviation, and medical care—range from $38 billion to $50 billion associated with the attacks, including the World Trade Center, the Pentagon in Washington, D.C., and Pennsylvania, making it the costliest U.S. disaster in the past two decades. Prior to the attacks, the largest insurance losses (in 2001 dollars) were the $19 billion damage caused by Hurricane Andrew in 1992 and the $14 billion damage caused by the Northridge, California, earthquake in 1994 (Schaad 2002).

13. Two widely cited reports were produced by the New York City Partnership and Chamber of Commerce (2001) and the New York City Office of the Comptroller (2001).

14. About 14 million square feet of space in the World Trade Center complex—World Trade Center Buildings 1, 2, 4, 5, 6, and 7—was destroyed. Estimates of the damaged commercial space in the World Trade Center area range from a low of about 14 million square feet, largely Class A space, to a high of 21 million square feet, which includes damaged Class B and C space. Estimates of the repair and replacement of the damaged commercial space are available for the Class A space only.

15. The estimates presented here are largely based on those reported by the New York City Partnership and Chamber of Commerce (2001), the New York City Office of the Comptroller (2001, 2002), and the Independent Budget Office (2002), updated with information that has become available since those studies were released.

16. See New York City Independent Budget Office (2002).

17. See New York City Office of the Comptroller (2001).

18. See New York City Office of the Comptroller (2002).

19. Estimates are based on a survey of large tenants (that is, occupying more than 10,000 square feet). See TenantWise (2002).

20. See New York City Independent Budget Office (2002).

21. The Bureau of Economic Analysis (BEA) of the U.S. Department of Commerce has estimated the property loss arising from the terrorist attacks on the World Trade Center and the Pentagon, treating the loss as a sharp increase in the depreciation of the fixed capital stock owned by private business and government. The value of the destroyed World Trade Center complex was based on its depreciated book value as opposed to replacement cost. The BEA estimates the total value of the assets destroyed in the attacks on the World Trade Center and the Pentagon at $15.5 billion. See U.S. Department of Commerce (2001).

22. Reliable welfare data for September, October, and November 2001 are not available.

23. The 1996 welfare reforms gave recipients incentives to move from the Temporary Assistance to Needy Families program to the Safety Net Assistance program. Examining the programs individually would confound the effects induced by these incentives with true changes in the rolls.

24. However, welfare caseloads may reflect a weakened economy with up to a two-year lag (Chernick and Reschovsky 2002).

25. These figures are based on data from appraisal firm Miller Samuel and calculations by the Federal Reserve Bank of New York.

26. These figures are based on data from the New York State Association of Realtors and the Office of Federal Housing Enterprise Oversight.

References

Asian American Federation of New York. 2002. "Chinatown after September 11: An Economic Impact Study." April 4.

Blank, Rebecca M. 2001. "Declining Caseloads/Increased Work: What Can We Conclude about the Effects of Welfare Reform?" Federal Reserve Bank of New York Economic Policy Review 7, no. 2 (September): 25-36.

Centers for Disease Control and Prevention. 2002. "Community Needs Assessment of Lower Manhattan Residents Following the World Trade Center Attacks—Manhattan, New York City, 2001." Morbidity and Mortality Weekly Report, no. 51 (special issue). September 11. <http://www.cdc.gov/mmwr/preview/mmwrhtml/mm51SPa4.htm> (November 2002).

Chernick, Howard, and Andrew Reschovsky. 2002. "Welfare Reform and Recession: Can the States Handle Both?" University of Wisconsin-Madison Institute for Research on Poverty FOCUS 22, no. 1: 22-8.

Dorman, Peter. 1996. Markets and Mortality: Economics, Dangerous Work, and the Value of Human Life. Cambridge: Cambridge University Press.

Fiscal Policy Institute. 2002. "World Trade Center Job Impacts Take a Heavy Toll on Low-Wage Workers." November 5. Available at <http://www.fiscalpolicy.org>.

Galea, Sandro, Jennifer Ahern, Heidi Resnick, Dean Kilpatrick, Michael Bucuvalas, Joel Gold, and David Vlahov. 2002. "Psychological Sequelae of the September 11 Terrorist Attacks in New York City." New England Journal of Medicine 346, no. 13: 982-7.

Mrozek, Janusz, and Laura Taylor. 2002. "What Determines the Value of Life? A Meta Analysis." Journal of Policy Analysis and Management 21, no. 2: 253-70.

New York City Independent Budget Office. 2002. "World Trade Center Aid: Too Much for Some Needs, Not Enough for Others?" Newsfax no. 100, May 28.

New York City Office of the Comptroller. 2001. "The Impact of the September 11 Attack on New York City's Economy and Revenues." October 4.

———. 2002. "One Year Later: The Fiscal Impact of 9/11 on New York City." September 4.

New York City Partnership and Chamber of Commerce. 2001. "Economic Impact Analysis of the September 11th Attack on New York City." November.

Orr, James, Robert Rich, and Rae Rosen. 1999. "Two New Indexes Offer a Broad View of Economic Activity in the New York–New Jersey Region." Federal Reserve Bank of New York Current Issues in Economics and Finance 5, no. 14 (October).

Schaad, Werner. 2002. "Terrorism—Dealing with the New Spectre." Swiss Re Focus Report, February.

Siena College. 2002. "Siena Research Institute: Surveys." <http://www.siena.edu/sri/> (September 2002).

TenantWise. 2002. "Special Report: Overview of Current Situation." July. Available at <http://www.tenantwise.com/072002wtc.asp>.

United Hospital Fund. 2002. "Disaster Relief Medicaid Offers Policy Lessons for the Future." Currents: Medicaid Managed Care 6, no. 4 (spring): 1-3.

U.S. Department of Commerce, Bureau of Economic Analysis. 2001. "Survey of Current Business." November.

Vlahov, David, Sandro Galea, Heidi Resnick, Jennifer Ahern, Joseph A. Boscarino, Michael Bucuvalas, Joel Gold, and Dean Kilpatrick. 2002. "Increased Use of Cigarettes, Alcohol, and Marijuana among Manhattan, New York, Residents after the September 11th Terrorist Attacks." American Journal of Epidemiology 155, no. 11 (June): 988-96.

James J. McAndrews and Simon M. Potter

Liquidity Effects of the Events of September 11, 2001

- On September 11, banks experienced difficulties in making their payments because of widespread damage to property and communications systems in Lower Manhattan. As a result, other banks received fewer payments than were expected.
- Since banks rely heavily on incoming funds to make their own payments, the normal coordination of payments broke down, and liquidity shortages developed at many banks.
- The Federal Reserve responded by supplying abundant liquidity to the banking system through discount window loans and open market operations—actions that helped restore payments coordination.
- The episode highlighted the usefulness of the discount window and of intraday lending by the central bank as tools for managing marketwide demands for liquidity.

In the wake of the terrorist attacks of September 11, 2001, the Federal Reserve supplied funds to the banking system in unprecedented amounts. The destructive force of the attacks themselves caused severe disruptions to the U.S. banking system, particularly in banks' abilities to send payments. The physical disruptions caused by the attacks included outages of telephone switching equipment in Lower Manhattan's financial district, impaired records processing and communications systems at individual banks, the evacuation of buildings that were the sites for the payments operations of large banks, and the suspended delivery of checks by air couriers.

These disruptions left some banks unable to execute payments to other banks through the Federal Reserve System's large-value payments system, Fedwire, which in turn resulted in an unexpected shortfall for other banks. Banks rely heavily on incoming funds to make their payments, so these unexpected shortfalls affected the distribution of balances across the banking system. The disruptions to the communications infrastructure also made it harder for banks to redistribute balances across the banking system in a timely manner. Accordingly, the actions of the Federal Reserve System were intended to counteract the effects of the unusual distribution of liquidity and the difficulties experienced by the banking system in distributing liquidity directly.

James J. McAndrews is a vice president and Simon M. Potter a senior economist at the Federal Reserve Bank of New York.

The authors thank participants at seminars at the Federal Reserve Bank of New York, the Federal Reserve Bank of Chicago, and the European Central Bank; Spence Hilton, Sandra Krieger, and Lawrence Sweet of the Federal Reserve Bank of New York; William Roberds of the Federal Reserve Bank of Atlanta; and Philip Strahan of Boston College for helpful comments. They also thank two anonymous referees, Daniel Bennett and Kurt Johnson for excellent research assistance, and the Wholesale Payments Product Office and the Credit Risk Management Function of the Federal Reserve Bank of New York for providing data for this article. The views expressed are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System.

The direct effects of the unusual distribution of balances and the difficulty of redistributing them resulted in a less regular flow of payments by banks to other banks. A disrupted and thus less coordinated payments flow can be thought of as a sudden drop in the velocity of high-powered money—that is, balances at the Federal Reserve. In the U.S. fractional reserve banking system, banks in aggregate regularly make payments that exceed their deposits at the Federal Reserve Banks by a factor of more than 100. To attain such a highly efficient utilization of balances requires a robust interbank money market as well as intricately harmonized timing and regularity in banks' payments activity. The sudden disorder in the regular timing of incoming payments made planning for a bank's liquidity needs more difficult. While some banks that experienced technological difficulties in sending payments accumulated higher-than-desired balances, other banks' increased uncertainty (regarding which payments they might receive later in the day) led them to have higher precautionary demand for liquid balances. Consequently, the sources of liquidity internal to the banking system were not available or capable of addressing the widespread demand for liquidity.

The demand for such large amounts of liquidity by U.S. banks in this episode might seem somewhat puzzling at first glance. Typically, students of banking history point to bank runs as examples of extraordinary demand for liquidity. However, there were no runs on banks in the United States following September 11. Nor were there extraordinary demands for currency by banks or by the public. Furthermore, the incident that triggered the liquidity shortfall was well known to all market participants and was not generally believed to cause any bank's credit quality or solvency to deteriorate significantly. Instead, the events drew further attention to the high level of interdependencies present in payments flows.

Given these interdependencies, the provision of liquidity by the Federal Reserve System allowed banks to make payments while being assured that funds would be available at the end of the day. This assurance helped banks resume their normal payments activities, which in turn increased the circulation of liquidity through the normal channels. In addition, in the aftermath of the attacks, banks and their customers engaged in extraordinary cooperative efforts to overcome the problems caused by the destruction of infrastructure and its attendant effects. (A review of many of these efforts is provided in Federal Reserve Bank of New York [2002].)

In this article, we provide new evidence of the importance of the coordination mechanism that banks use in their normal payments activity to provide liquidity. We do so by first outlining the sources of funding for banks' payments activity, thus highlighting the role of expected incoming payments from other banks. Next, we examine the disruptions in the payments and communications mechanisms on September 11 and the disorder that resulted from those disruptions. We then directly estimate banks' payments reactions to the receipt of payments from other banks and how their reactions changed in the days following the attacks. Finally, we discuss possible alternative ways to avoid such large disruptions to the payments mechanism, as well as ways to settle payments and the potential of these ways to reduce liquidity demands in the event of a significant disruption to the payments mechanism.

Sources of Liquidity for Payments: The Strategic Dimension

Banks must have enough liquidity to cover the amount of their payments made over Fedwire.[1] This liquidity typically consists of: 1) balances maintained on account with the central bank, 2) borrowing from other banks through the money markets, 3) credit extensions from the central bank, and 4) expected incoming transfers from other banks.[2] The last category creates a strategic challenge for banks in their liquidity management: they each rely on one another's payments as a source of funding.

The first source of funds available for making a payment is the balances kept on account at Federal Reserve Banks. For commercial banks, these balances consist of either required reserve balances, excess reserve balances, or service-related balances.[3] These balances and service-related balances for August 2001 averaged $14.65 billion per day.[4] Banks are required to maintain these balances at a certain level during two-week periods known as reserve maintenance periods. In addition, the Federal Reserve can supply funds to the banking market through open market operations. By purchasing securities, the Federal Reserve directly increases banks' balances held on account at System Banks.

Overnight balances at the Federal Reserve are costly to maintain because they do not earn interest.[5] Nevertheless, if banks' balances fall below the target on average for the two-week period, the banks face a penalty rate and must hold a higher level of balances during the next two-week period. In addition, if banks fall into overdraft positions on any given night, they must pay a substantial penalty of 4 percentage points in excess of the effective federal funds rate for that day. As a result of these disincentives to falling short of required balances and to holding excessive balances, banks try to target their overnight balances within a narrow band.

If a bank's payments over the course of the day exceed its receipts by more than the value of the balances with which it started the day, a bank can borrow funds in the federal funds market. Federal funds activity averaged $144 billion per day in the first quarter of 1998 (Furfine 1999).[6] Fed funds activity redistributes the liquidity among banks but does not add to the approximately $15 billion in deposits in the Federal Reserve System. The Federal Reserve can increase the deposit level by lending funds through the discount window,[7] where borrowing averaged only $174 million per day in August 2001.[8] Using the fed funds market and the discount window as sources of funding comes at a cost. Fed funds purchases are uncollateralized borrowing from other commercial banks, and the borrower pays the going fed funds rate to compensate the lender for the risk involved. Loans from the discount window are made at the discount rate, which is generally lower than the fed funds target rate. However, banks are not expected to rely on the discount window for funds on a regular basis.[9]

Intraday credit from the Federal Reserve System provides another source of funding for banks' payments.[10] Coleman (2002) reports that the average value of daylight overdrafts, as this form of intraday credit is called, was $32.8 billion per day in August 2001. However, banks tend to economize on this source of funding because the Federal Reserve System charges banks a fee for daylight overdrafts above a certain amount.

In general, the daily value of banks' payments greatly exceeds the value of those sources of liquidity. Consider that during August 2001, the value of Fedwire funds transfers averaged more than $1.6 trillion per day, while banks held about $15 billion on account.[11] In other words, the "turnover" of each dollar on banks' balances was more than 100 (if security transfers are included, the turnover ratio is about 180).[12] Banks can make payments even though the value of those payments is considerably higher than the value of the sources of funds because most payments are offset over the course of the day. Most banks can expect that they will receive a certain value of incoming payments during the day. To some extent, a bank can use these expected receipts to plan its submission of payments over Fedwire throughout the day. Of course, other banks are planning to use incoming funds to make payments as well, so how can banks in aggregate use expected receipts as a source of funding?

The Strategy of Payments Coordination

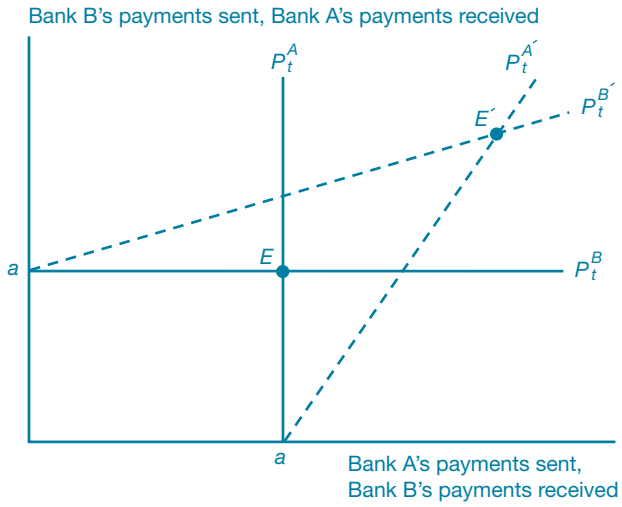
Banks use their scarce liquidity to serve both their customers' payments needs and to complete their own agreed-upon payments arising from trading activities and interbank lending. As we noted, banks attempt to make use of incoming funds received from other banks by strategically timing payments.[13] In models of this strategic payments behavior, banks make payments in reaction to their receipts from other banks. Each bank behaves by choosing the best reaction available to it, given its receipts. A bank's "reaction function" describes how a bank responds to payments received.

In general, we can identify at least two ways in which the receipt of a payment by a bank can increase the likelihood of that bank making a payment. First, the receipt of the funds replenishes the bank's balances at the central bank. Given that the bank has a list of payments it wishes to make during the day, the receipt of funds allows the bank to make an outgoing payment with less chance of incurring an overdraft in its account, all else being equal. In other words, some payments activity can be "self-funded" by matching outgoing payments to incoming payments. Second, as a payment is received on behalf of a customer, that customer is subsequently more creditworthy. If the customer had been at its credit limit prior to the receipt of payment, the receipt into its account relaxes the credit constraint of the customer. The customer's subsequent requests to the bank to make payments on its behalf will not be delayed as a result of the customer's credit exposure to the bank.

Exhibit 1 shows, for simplicity, the activity of a payments system in which only two banks participate. The horizontal axis shows payments sent by Bank A and the vertical axis shows payments received by Bank A, and vice versa for Bank B. Two sets of reaction functions are shown. A reaction function shows

Exhibit 1

Payments Reaction Curves and Equilibrium



the level of a bank's payments as a function of its expected receipts over some period of time. This function can be represented by the algebraic relationship $P_t^A = a + bR_t^A + \varepsilon_t$, where P_t^A is Bank A's payments in time t, R_t^A represents the receipts of Bank A at time t (both expressed in dollar values), a is the bank's autonomous willingness to send payments (irrespective of its receipts), and the parameter b represents the slope of the reaction function. ε_t is an error term at time t.

The solid-line reaction curves in the exhibit show the banks sending payments without considering their receipts. The reaction curves have a slope of zero with respect to the receipts from their counterparties (in the case of Bank A, its reaction curve is vertical with respect to the horizontal axis, but has a zero slope with respect to its receipts, which are shown on the vertical axis). The intercepts show the autonomous willingness of each bank to make payments and are functions of each bank's balances on deposit at the central bank. The equilibrium of payments occurs at the intersection of the two reaction curves, labeled *E*. In contrast, the dashed-line reaction curves have a positive slope with respect to the receipts from the banks' counterparties, reflecting "strategic complementarity."[14] For these reaction curves, the equilibrium occurs at the point labeled *E′*. The positively sloped reaction curves—or strategic complementarity—are significant because they show that in equilibrium, the banks can make more payments per dollar of balances. This is the case because at *E′*, banks conduct more payments starting from the same level of bank deposits at the central bank (the intercepts being the same for both sets of reaction functions). It is in this sense that payments are coordinated in equilibrium. The lower the level of coordination—or the closer the slopes of the reaction curves are to zero—the greater a bank's balances at the central bank must be for that bank to make a given value of payments. (Later in this article, we confirm that the slopes of banks' reaction functions fell significantly in the days after September 11, 2001.)

One way that the banking system can coordinate payments and use expected incoming funds as a source of liquidity (in a real-time gross settlement system) in equilibrium is by submitting payments simultaneously, or almost simultaneously.

One way that the banking system can coordinate payments and use expected incoming funds as a source of liquidity (in a real-time gross settlement system) in equilibrium is by submitting payments simultaneously, or almost simultaneously.[15,16] In the Fedwire funds transfer service, banks typically send the bulk of their daily payments between 3:30 p.m. and 5:30 p.m., and this pattern is very stable. In that way, banks recirculate liquidity very quickly and take advantage of offsetting payments without incurring daylight overdrafts for an extended period of time.

Another strategic element in making payments (in an RTGS system) is the risk to which a bank is subject when paying out funds in advance of receiving funds from a counterparty.[17] A perceived increase in this risk can result in banks delaying the submission of their payments while they await the arrival of payments from counterparties.[18] The actions of a central bank cannot directly reduce this risk, but they can weaken the spillover effect of this risk on banks' willingness to send payments to third-party banks. Under normal circumstances, this risk is relatively low, but in the days following September 11, the risk may have affected banks' willingness to submit payments in a timely manner, notwithstanding the pledge of the Federal Reserve System to provide low-cost liquidity to the banks.

Example of a Coordination Failure

If a bank fails to make an expected payment, the ability of other banks to make payments can be disrupted in that they send fewer payments, which results in fewer receipts by their counterparties, thus creating a downward cycle.

Consider the following example of three banks, each of which has to pay one of the other banks once during the day. Assume overdrafts are not allowed. Exhibit 2 illustrates the situation. Suppose that under normal circumstances, Bank A pays Bank B every morning. Bank B, using funds transferred from Bank A, pays Bank C; then Bank C can pay Bank A. However, if Bank B fails to pay Bank C, then Bank C cannot pay Bank A. Bank B accumulates deposits at the Federal Reserve as its "due froms" are paid to it by a Fedwire transfer, while Bank A and Bank C together have reduced deposits at the Federal Reserve but have not reduced their "due froms." Thus, the inability of a bank to send out payments causes that bank's deposits to pile up and reduces the ability of other banks to send funds.

One measure of concentration in banks' balances at the central bank is the Herfindahl-Hirschman index (HHI).[19]

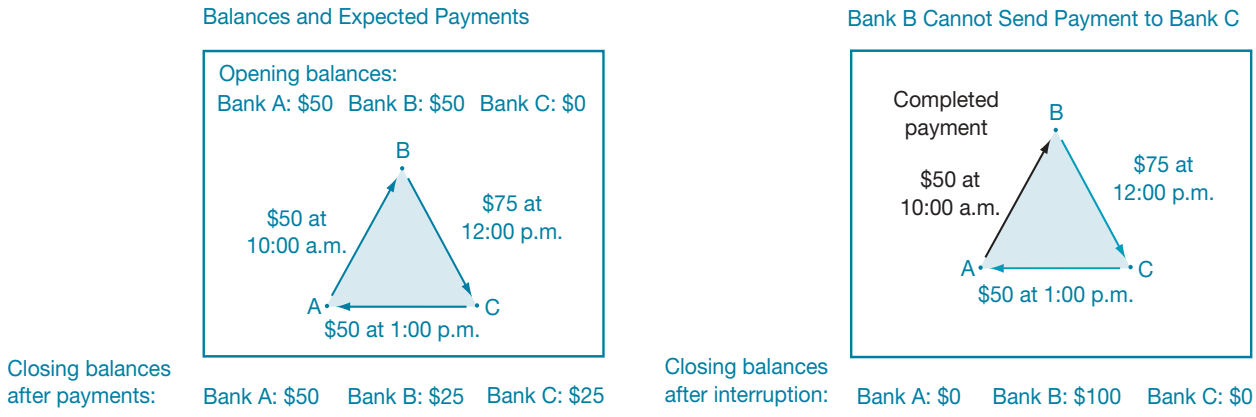
As shown in Exhibit 2, the concentration level in deposits at the central bank rises considerably when the expected payment from Bank B is not made (from either the opening level of balances or from expected end-of-day balances), leaving Bank A and Bank C unable to make payments.

A measure of liquidity usage is provided by the system's *turnover ratio*—the ratio of the value of total payments made to total deposits. In the exhibit, the turnover ratio falls from 1.75 to 0.5. The lower turnover ratio shows that the payments made are more costly in terms of bank balances, which indicates an increase in the demand for liquidity per dollar of payments.

Another measure of liquidity is the *netting ratio*. This is the ratio of the day's total payments to the amount of funds that would need to be transferred between accounts if only the net amounts flowing between banks were exchanged.[20] In the exhibit, the netting ratio falls from 7 to 1 as the expected payment—and subsequent payments—fail to be made. The expected level of 7 shows that the banks would only need to hold $25 (to be held by Bank B) if they were to make all payments simultaneously (or transfer only net amounts) and still comply with the rules that their balances not fall below $0. The actual level of the netting achieved shows again that the

Exhibit 2

Example of the Effects of One Bank's Failure to Send Payments



Assets and Liabilities of Banks at Opening of Day and after Bank B Fails to Pay Bank C

	Bank A		Bank B		Bank C	
	Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
Opening	Balances at central bank: $50	Due to Bank B: $50	Balances at central bank: $50	Due to Bank C: $75	Balances at central bank: $0	Due to Bank A: $50
	Due from Bank C: $50		Due from Bank A: $50		Due from Bank B: $75	
	Other	Other	Other	Other	Other	Other
After interruption of payments	Balances at central bank: $0	Due to Bank B: $0	Balances at central bank: $100	Due to Bank C: $75	Balances at central bank: $0	Due to Bank A: $50
	Due from Bank C: $50				Due from Bank B: $75	
	Other	Other	Other	Other	Other	Other

actual payments are more costly than expected and—to the extent that banks wish to make payments—indicates that demand for liquidity rises.[21]

Bank C, if it still wishes to make the payment to Bank A (or if there are more "rounds" of payments to be made later in the day, or on the next day), can borrow funds on the interbank market or borrow from some other party, typically the central bank. In addition, the central bank can add reserves to the system through open market operations.

If it were to borrow on the interbank market, from whom would Bank C borrow? Only Bank B has funds to lend. It might be the case that although Bank B cannot feasibly send the expected payment to Bank C, perhaps because communications with its customer are impaired, Bank B is able to lend to Bank C. In this case, there is an increase in demand for borrowing on the interbank market, and Bank B would lend more than usual in that market. The payment by Bank C to Bank A would be accomplished. If Bank B cannot easily make an interbank loan because of technical difficulties, then the central bank can inject reserves either through the discount window or open market operations. In either case, Bank C would make its payment, but, once again, the turnover ratio would fall. In this case, the turnover ratio falls not necessarily because payments values fall, but rather because the amount of reserves in the system needed to accomplish payment has risen.

The example illustrates the fragile nature of using payments coordination as a source of liquidity for banks. One bank's failure to pay as expected because of a technological problem can be felt as a liquidity shortage by the bank's counterparties, and their counterparties, making payments activity more costly.

Payments Disruptions Following September 11, 2001

After the attacks on the morning of September 11, it was immediately clear to financial market participants that general operations and communications and computer systems in Lower Manhattan were not functioning well.[22] A direct effect of these difficulties was a reduction in the value and volume of transfers on Fedwire on September 11.[23] Although there were major disruptions to markets during that week—the New York Stock Exchange was closed for four days and the commercial paper market was significantly disrupted—it is worth noting that settlements occurred at the major large-value private-sector settlement systems (the Depository Trust & Clearing Corporation and the Clearing House Inter-bank Payments System [CHIPS]) on the eleventh and subsequent days.

Information on aggregate activity for the banking system on September 11 and the days surrounding it is presented in Table 1. The physical disruptions caused by the attacks are evident in the sharp drop in the volume of payments made on Fedwire. Furthermore, one can observe that the volume of payments remained low for the next two days as physical disruptions caused by the attacks continued to affect banks and their customers.

The value of funds sent on September 11 was $1.2 trillion, about three-fourths of the average for the benchmark period. However, unlike volume, the value of funds sent had returned to normal levels on the twelfth and was then at elevated levels for the next seven business days.

The aggregate balances of the banking system increased starting on September 12 and reached a peak of $121 billion on September 14. By Friday, September 21, balances returned to normal levels. Although the Federal Reserve System injected funds into the banking system in a number of ways, the pattern of an increase in balances is mainly explained by discount window loans and open market operations. Recall the turnover ratio, which is the ratio of the value of total payments made to total deposits. In our example, the turnover ratio fell as liquidity was not distributed as expected via the payments system. As we see in Table 1, the turnover ratio for Fedwire funds transfers fell from more than 100 before September 11 to only 18 on September 14, indicating a significant increase in the amount of liquidity used to make a dollar's worth of payment. Once again, by the end of the next week, the turnover ratio was at normal levels.

Table 1

Fedwire Funds Transfer Value and Volume, and Aggregate Opening Balances with the Federal Reserve: September 10-21, 2001

Date	Volume	Value (Billions of Dollars)	Balance (Billions of Dollars)
September 10	436,312	1,591	13
September 11	249,472	1,216	13
September 12	332,433	1,696	44
September 13	376,937	1,952	104
September 14	423,256	2,009	121
September 17	462,522	2,312	111
September 18	419,126	1,978	46
September 19	401,420	1,836	19
September 20	433,771	1,921	15
September 21	442,293	1,832	13

Source: Federal Reserve Bank of New York.

We also examined two alternate estimates of fed funds activity during the period.[24] These estimates indicate that some banks were unable to return funds on September 11 that they had borrowed on September 10, presumably because of the physical disruptions following the attacks. Furthermore, because of the physical disruptions, a smaller than average amount of fed funds was sent on September 11. For the next few days, however, fed funds activity was elevated as the Federal Reserve System implemented large open market operations.

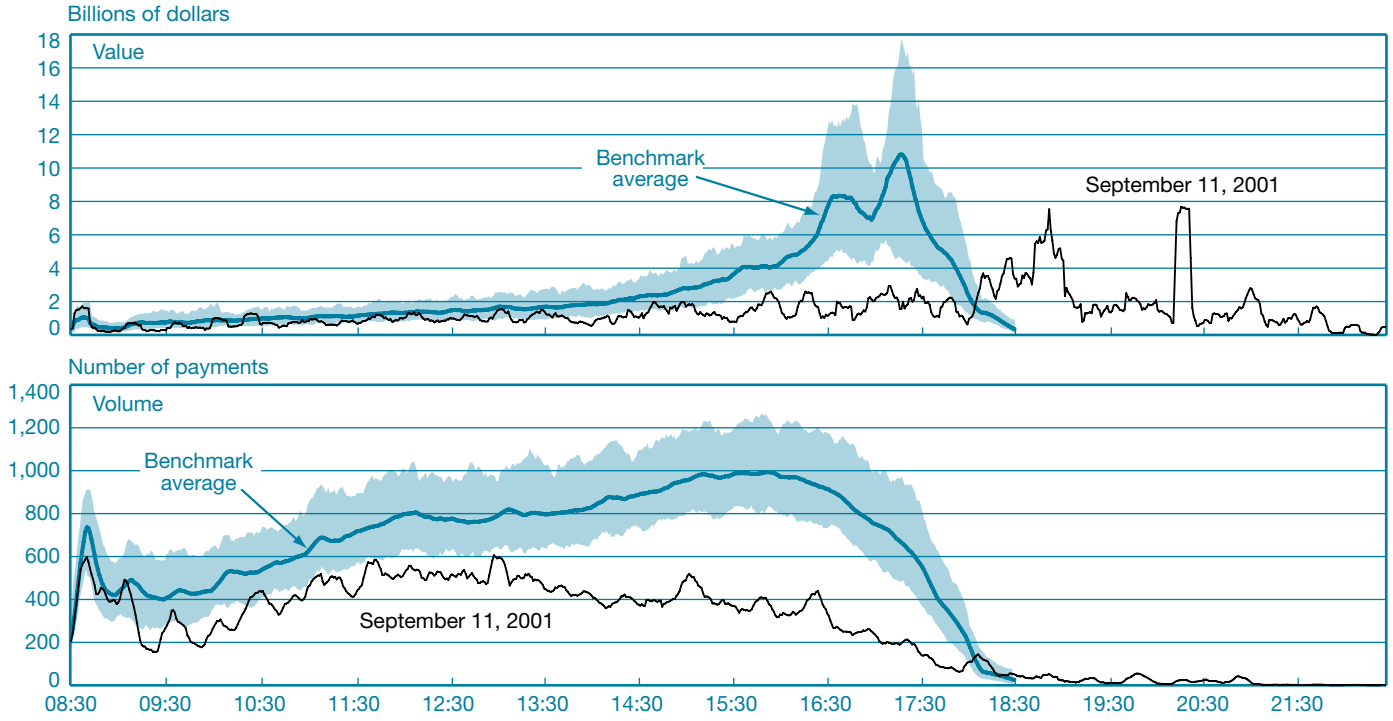
Although the values of Fedwire funds transfers and fed funds activity were quickly restored, the timing pattern of these payments was disturbed for a more extended period. As the day of September 11 progressed, it became clear to banks not directly affected by the problems in Lower Manhattan that both the level and timing of activity on Fedwire were significantly affected. Furthermore, banks that were directly affected were having difficulty sending out payments, communicating with customers, maintaining up-to-date records, and delivering securities.

Chart 1 compares the timing of payments on September 11 with the average for the benchmark period (the benchmark period serves as our measure of normal payments activity and allows us to gauge the effects of the attacks).[25] The top panel shows the value transferred per minute and the bottom panel shows the volume of payments transferred per minute. Both panels include a two-standard-deviation band around the benchmark period average to indicate the normal range of variability. As can be seen in Chart 1, another direct effect of the disruptions was a very different pattern of payments timing compared with the benchmark pattern. In particular, more than one-third of the value of payments was sent after the usual closing of Fedwire at 6:30 p.m.

Chart 1

Value and Volume of Funds Sent per Minute

Ten-Minute Moving Average



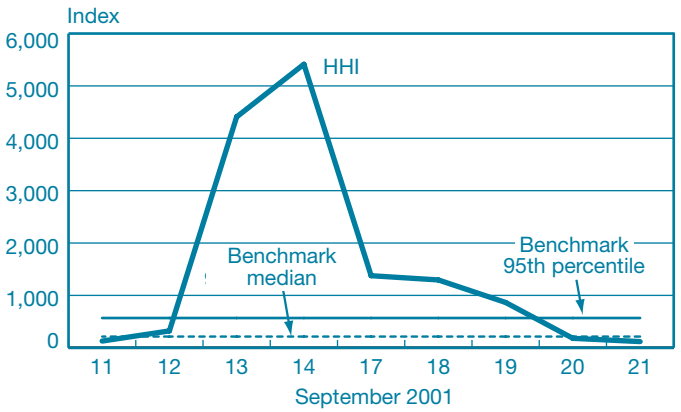
Source: Federal Reserve Bank of New York.

Note: The shaded bands indicate +/- two standard deviations of the benchmark averages.

Concentration in Bank Balances

Chart 2 shows the median and 95th percentile of the HHI of account balances (for opening balances) in the benchmark data set and for each of nine business days in the attack data set.[26] It is clear that the concentration in account balances at the Federal Reserve—rising more than fourteen-fold from its normal levels on the days following the terrorist attacks—was a most unusual event. Furthermore, as we will show, this measure would be even higher if the Federal Reserve had not distributed funds from the discount window in the pattern it did.

Chart 2
Herfindahl-Hirschman Index for Start-of-Day Federal Reserve Balances, by Day

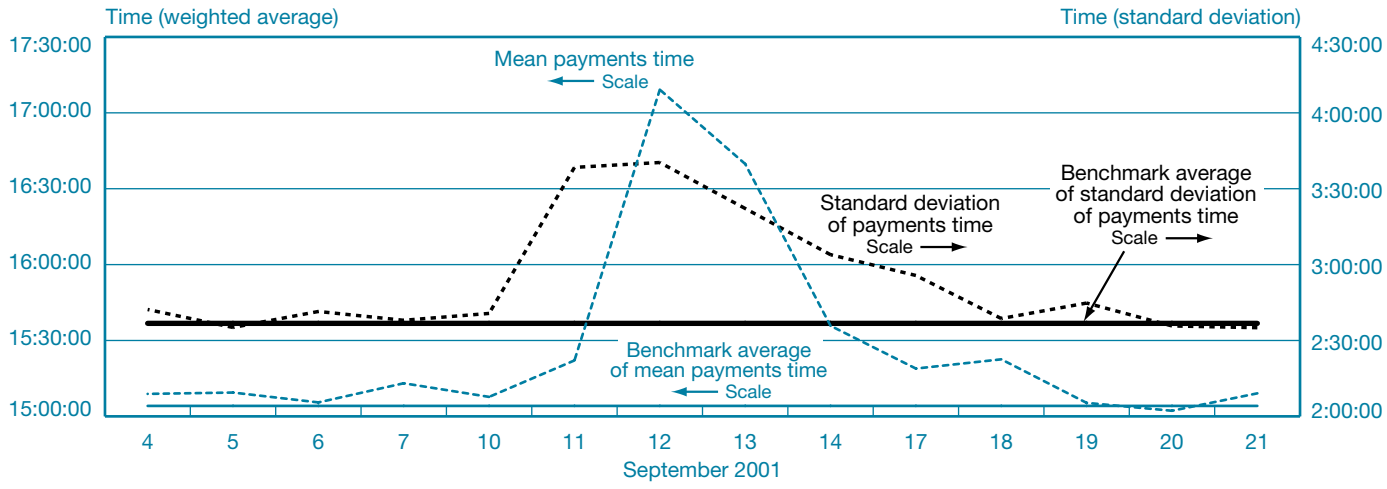


Source: Federal Reserve Bank of New York.

Measures of Coordination Failures

Next, we examine two measures of the tendency of payments activity to be concentrated in a particular time during the day. Chart 3 shows the standard deviation of the timing of payments weighted by their value and the average time of payments, also weighted by value. The standard deviation measure indicates that the variability of payments timing rose precipitously—almost doubling to about four hours from two—indicating an unusual and disordered pattern of timing of payments relative to the benchmark period. In addition, payments were significantly delayed. This is shown by the value-weighted average of the time of payment on Fedwire, which was delayed by approximately two hours relative to its benchmark-period average.[27] The same patterns are present in the estimates of fed funds activity as well; the borrowing and

Chart 3
Value-Weighted Average of the Timing of Payments on Fedwire
Benchmark Period and September 4-21, 2001



Source: Federal Reserve Bank of New York.

lending in the banking system's internal market for distributing liquidity occurred later in the day and with a much more varied pattern during the remainder of the week. Both of these changes are highly significant relative to normal patterns.

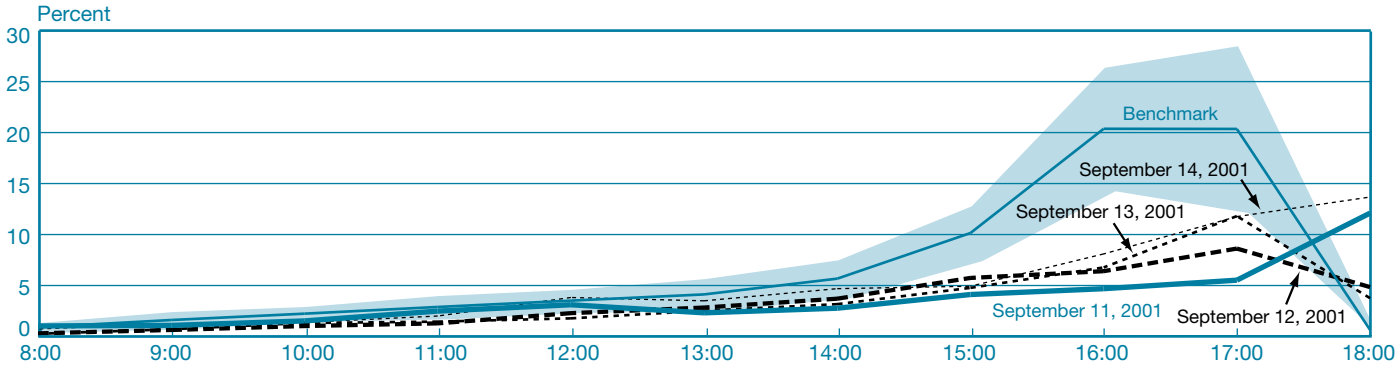
Another indication of the disorder in payments during those days was the decreased netting ratio of payments. Recall that the netting ratio is the ratio of the day's total payments to the amount of funds that would need to be transferred between accounts if only the net amounts flowing between banks were exchanged. Chart 4 displays a related measure of netting. It shows the amounts of payments in terms of the percentage of the day's total payments offset within each hour of the day. We shaded the area on the chart that represents the mean netting ratio plus or minus two standard deviations for each hour across the 107 days in the benchmark data set. It is clear that the fall in the concentration of payments was accompanied by a drop in the degree to which banks managed to arrange offsetting payments to be made at roughly the same time, particularly in the late afternoon period. Between 3:30 p.m. and 5:30 p.m., banks regularly offset about 20 percent of the day's total payments per hour, and the values transferred per minute are usually the greatest. On the four days immediately following the attacks, offsetting payments made during those two hours fell to between 5 percent and 10 percent of the day's total payments per hour. A day's payments activity averaged $1.685 trillion in 2001, so a decline of 10 percentage points in offsetting payments during those two hours could result in an increased demand for liquidity of $163 billion, all else being equal. The coordination of payments timing is crucial in reducing the banks' demand for liquidity. The value of intraday overdrafts on September 11-13 was much higher than usual because fewer offsetting payments were made.[28]

Payment Reaction Function Estimates

We estimate the reaction function of bank payments over Fedwire by measuring a bank's own tendency to make outgoing payments as a function of its receipts. In the equation, we posit that a bank's decision to send outgoing payments is dependent on the payments it receives from all other banks, as in the equation we described earlier, $P_t^A = a + bR_t^A + \varepsilon_t$ (this formulation is similar to that of Bech and Garratt [forthcoming], who consider a two-agent game).[29] By focusing on the slope coefficient b, we can judge how much strategic complementarity, shown by a positive slope, is present in banks' payments timing strategies.

We estimate the reaction function by pooling the Fedwire activity of twenty large banks—whose payments activity accounts for more than 60 percent of Fedwire volume—together in a panel estimation with fixed effects (Table 2). In this panel approach to estimating the reaction function, a bank's payments in a one-minute interval for each minute during the day are dependent on the bank's receipts in the previous fifteen-minute interval, the bank's opening balance, and the bank's cumulative receipts minus its cumulative

Chart 4
Percentage of the Day's Total Payments Netted during Each Hour



Source: Federal Reserve Bank of New York.

Note: The shaded band indicates +/- two standard deviations of the benchmark averages.

payments sent up to sixteen minutes prior to the minute.[30] In addition, we include dummy variables for the time period before 8:30 a.m., for the period between 6:00 p.m. and 6:30 p.m., and for the tenth day of a reserve maintenance period.[31] We perform another estimation including dummy variables for each half hour of the day between 2:00 p.m. and 6:00 p.m. These latter time-of-day dummies are intended to capture any institutional regularity or very long-run patterns of coordination that are common across banks in their daily payments behavior that might contribute to the pronounced peak in payments in the afternoon (Table 2, panel A). We also include a bank-specific intercept (fixed effect) for each bank in both estimating equations.

The parameters of this equation are estimated under a Tobit assumption for the functional form. That is, we use the same equation, first to predict if a payment would be made in the relevant time interval, and then to predict the payment's size if a payment is to be made. The results from estimating the coefficient on a bank's receipts separately for each of the five days after the attacks are displayed in Table 2. In addition, Table 2 shows the results obtained from pooling the data over the entire benchmark period.[32] This coefficient on receipts is a good estimate of the reaction function's slope for large payments, and we label the coefficient estimate the slope. The slope of the reaction function represents a bank's marginal propensity to send out payments in response to the receipt of payments from its counterparties.

For the benchmark period, the slope is positive, as expected, and has a very precisely estimated value of 0.765. This relatively steep slope implies that banks have achieved a high degree of coordination in their payments activities. At the same time, the steep slope of the reaction function displays strong strategic complementarity among the banks' desired actions, implying potentially large effects of a breakdown in coordination. Adding the time dummies for the afternoon period—to capture institutional or long-run patterns of payments behavior—reduces the size of the coefficient marginally to 0.632 (Table 2).

In contrast, the reaction function slope drops considerably for the first four days after the attacks. The slope is estimated precisely enough for each day that the decline is statistically significant: the estimates on those days are lower than the benchmark estimate. We also estimated the coefficients individually for the five business days before September 11 to assess the variability in the reaction function slope within the benchmark period. Chart 5 plots these results along with the estimates from the nine business days after the attacks. This shows that the average from the benchmark period accurately captures the daily behavior in the days before the attacks. Furthermore, in the week after the attacks, we find that the slope of the reaction function increased above the benchmark level.

The pooling of data across banks assumes symmetry of reaction functions. We also estimated individual bank reaction functions. The results of these estimations are very similar to those of the pooled estimation. The weighted average of the slopes of the reaction is similar in size and follows the same pattern of a sharp dip below the benchmark estimate in the four days following the attacks and a move above the benchmark estimate in the week of September 17. These

Table 2
Reaction Function Slopes

	Benchmark	September 11, 2001	September 12, 2001	September 13, 2001	September 14, 2001	September 17, 2001
	Panel A: Estimates of the Coefficient on Receipts of the Pooled Model					
Reaction function slope	0.765	0.229	0.358	0.199	0.599	1.190
	(0.0023)	(0.0528)	(0.0379)	(0.0419)	(0.0394)	(0.0312)
	Panel B: Estimates of the Coefficient on Receipts of the Pooled Model Including Afternoon Time Dummies					
Reaction function slope	0.632	0.202	0.249	0.083	0.466	0.863
	(0.0027)	(0.0534)	(0.0389)	(0.0455)	(0.0402)	(0.0343)

Sources: Wholesale Payments Product Office; Federal Reserve Bank of New York Credit Risk Management Function.

Notes: The number of observations is 2,259,920 for the benchmark and 27,920 for each day listed. Standard errors are in parentheses.

individual-bank estimates yield another insight, namely, that the standard deviation of the estimated slope of the reaction function across banks was significantly elevated throughout the period of September 11-14, indicating that banks' actions were highly uncoordinated during that whole period and only began to return to a coordinated set of reactions on September 17.

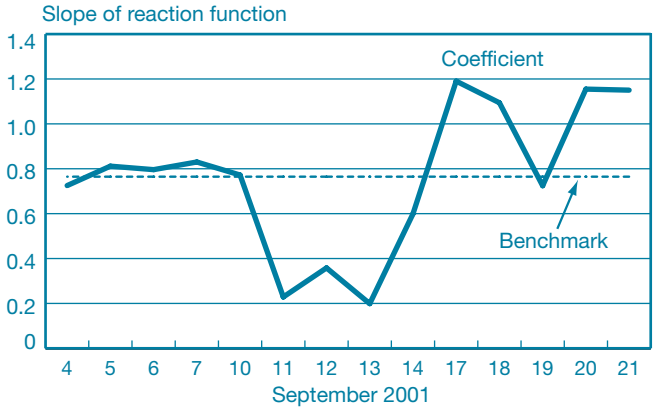
These results imply that the coordinating equilibrium in the payments flows on Fedwire was greatly disrupted by the events of September 11. In particular, we estimate that the slope of the reaction function fell significantly in the four days immediately after the attacks, indicating that to make a given value of payments, banks held higher balances at the Federal Reserve. This shows the tendency of localized problems in the payments system to spill over to the whole system by the presence of strategic complementarity.

There are at least two interpretations of these results. One is that the increase in the disorder of payments receipts by banks caused a change in the expectations of those banks to receive funds in the normal pattern, and so they experienced an increase in precautionary demands for balances. Consistent with this interpretation, payments activity occurred later in the day and in a more variable pattern than usual, and banks did hold much higher balances. Furthermore, although the communications infrastructure problems were overcome gradually and were largely resolved by Friday the fourteenth, the disorder in payments persisted on Friday and the high balances were observed for several additional days. Another interpretation is that the physical destruction of the communications and business infrastructure made the usual process of payments coordination more difficult—a situation that would heighten the demand for precautionary balances. For example, if a customer requests a payment to be made on its behalf, but it is at its credit limit, a bank will often check with counterparties to determine if that customer is to receive a payment from a counterparty. In addition, internal communication within the bank takes place among credit officers and payments personnel to ensure that approvals are in place prior to the release of a payment. As communication was disrupted generally, these processes occurred more slowly and with greater difficulty. Such obstacles to making payments in a timely fashion would tend to result in a decrease in the slope of the reaction function.

[Our results show] the tendency of localized problems in the payments system to spill over to the whole system by the presence of strategic complementarity.

The reaction function estimates for the five business days starting on September 17 indicate that the system quickly managed to regain a coordinating equilibrium. How did the system manage to reestablish the coordinating equilibrium? One direct impetus, of course, was the resolution of technological problems at individual banks.[33] However, this does not necessarily explain the timing of the increase in the reaction function slope estimate, or its higher level (compared with the benchmark) in the week of September 17-21. For a possible explanation of this pattern, we now turn to the actions of the Federal Reserve System.

Chart 5
Slope of Reaction Function of Payments Sent to Payments Received: Fixed-Effects Tobit Model



Source: Federal Reserve Bank of New York.

The Actions of the Federal Reserve System

The Federal Reserve System took a number of steps to address the problems in the payments system after September 11, 2001. Around noon on the eleventh, the Board of Governors of the Federal Reserve System released a statement saying, "the Federal Reserve System is open and operating. The discount window is available to meet liquidity needs."[34] In addition, for the period from Tuesday, September 11, through Friday, September 21, it waived daylight overdraft fees and overnight

overdraft penalties.[35] (These actions significantly increased banks' account balances at the Federal Reserve, as shown in Table 1.) Coleman (2002) provides a useful summary of the Federal Reserve's actions for those four days:

> "Peak and average daylight overdrafts that depository institutions incurred were approximately 36 and 32 percent higher, respectively, than levels in August 2001" (p. 81).
>
> "Overnight overdrafts increased from an average of $9 million in August 2001 to more than $4 billion on September 12 (and the penalty fee was waived on overnight overdrafts). Discount window loans rose from around $200 million to about $45 billion on September 12; later, when markets began to function better, Federal Reserve open market operations increased from $25 billion to nearly $100 billion" (p. 82).

Federal Reserve staff also contacted banks often during September 11-13, encouraging them to make payments and to consider use of the discount window to cover unexpected shortfalls that the banks might encounter later in the day. The assurance provided by the Federal Reserve's press release and the statements of discount window officials to banks may have prevented a difficult situation from becoming worse.

Once banks' demands for liquidity had been met at low cost and their need to rely on incoming payments had been reduced, banks could once again send payments in more normal patterns. The return to an earlier average payments time on September 14 (Chart 3) indicates that banks were sending payments more freely early in the day compared with the previous two days. Our reaction function slope estimates also point to a return to more strategic complementarity in payments activity on the fourteenth.

In the following week of September 17-21, the Federal Reserve, through open market operations, maintained significant additional liquidity in banks' accounts, with the fed funds rate trading below the target interest rate. During that week, the data suggest that coordination reached higher-than-normal levels, as measured by the steepened slopes of the reaction functions. This is likely a consequence of the low cost of liquidity from overnight balances at the Federal Reserve. Conditions for greater-than-normal flows were present with a low cost of overnight balances, no lingering credit problems at banks, and a backlog of payments at banks to be processed.

Our results, combined with the actions of the Federal Reserve System, can be understood with the assistance of Chart 6. The upper left panel shows the normal (benchmark period) reaction functions of two representative banks, Bank A and Bank B (using hypothetical data that are consistent with the data and estimates of our model). The intercept of the reaction function represents a bank's autonomous willingness to send payments, even if it receives no payments from other banks. The intercept also depends on a bank's opening balances. The upward slope of the reaction function is consistent with the strategic complementarity present in the reaction functions we estimated.

Once banks' demands for liquidity had been met at low cost and their need to rely on incoming payments had been reduced, banks could once again send payments in more normal patterns.

The upper left panel shows a second set of reaction functions illustrating the effect of an interruption of Bank B's ability to send payments in response to receiving a payment from Bank A. In this set, Bank B is unable to send out payments until the end of the day. As shown in the chart, Bank B has received many payments, leading to a high balance in its account. Bank A, in contrast, has a low balance and a reduced autonomous willingness to send payments.

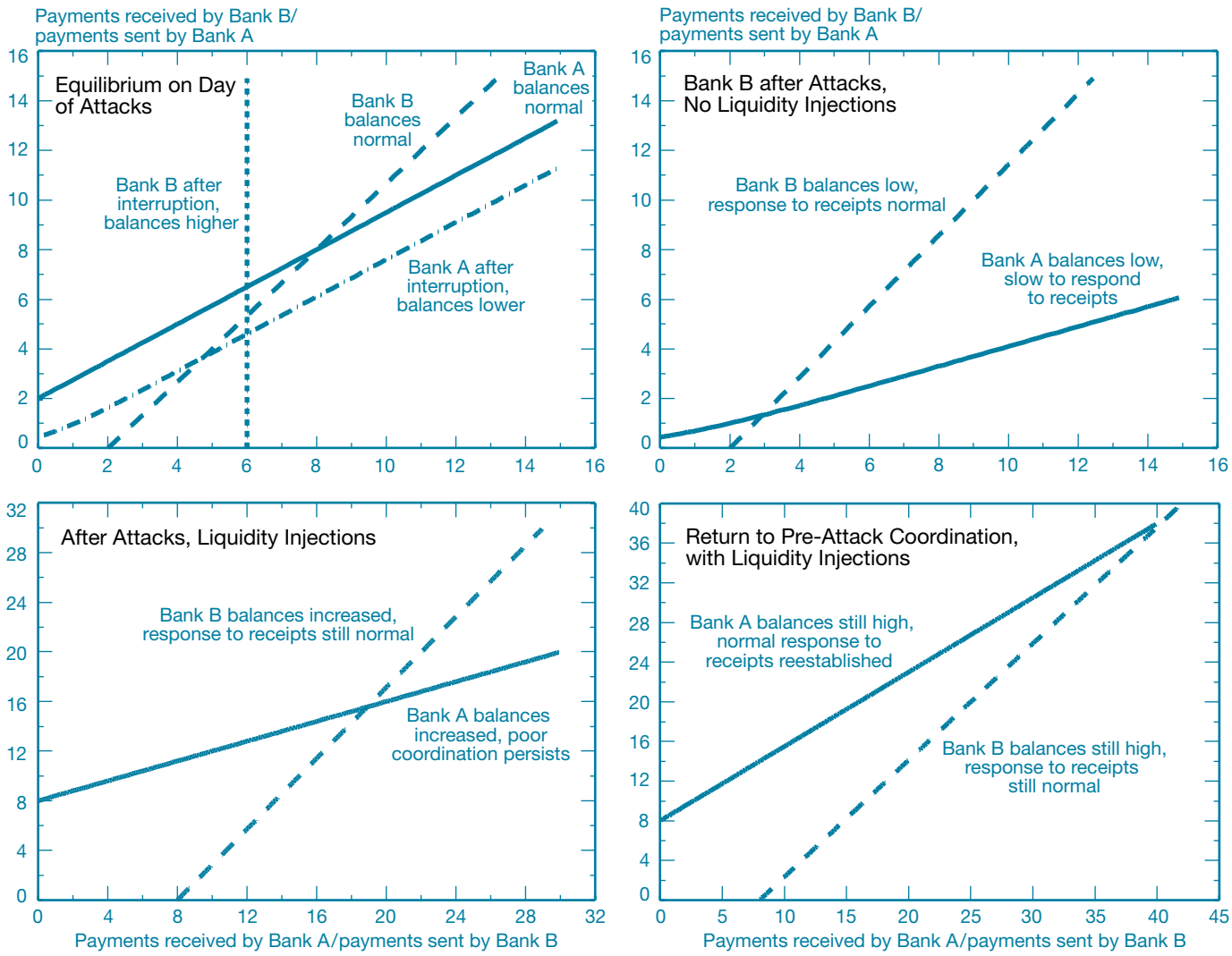
The upper right panel shows the set of reaction functions immediately following Bank B's recovered ability to respond to Bank A. In response to Bank B's previous inability to react to a payment, Bank A has lowered its expectations of receiving further payments. This lowers Bank A's willingness to send out payments in a coordinated fashion. Thus, payments activity is at a low level despite the relatively steep slope of Bank B's reaction function.

The lower left panel shows the effects of liquidity injections by the Federal Reserve System. The increase in balances raises the banks' autonomous willingness to send payments and results in an increase in payments value and in turnover.

Finally, in the lower right panel, the banks' coordination increases further while the liquidity supplied by the Federal Reserve remains in the System. This panel represents the situation during the week of September 17-21. The combined effects led to higher values of payments and payments coordination, as observed in the Fedwire activity that week.

Chart 6

Reaction Functions Illustrating Payments Coordination after the Attacks



Source: Authors' calculations.

Discussion

The Federal Reserve Response

From the point of view of the strategic models of payments timing and submission discussed earlier, the actions of the Federal Reserve System to provide liquidity to the banking system following the attacks of September 11 were appropriate. The Federal Reserve acted to provide liquidity to banks—liquidity that many banks would have expected to receive under more normal circumstances in the form of incoming payments or via fed funds borrowing.

But precisely how effective were the actions of the Federal Reserve in restarting the coordination of payments activity? That question is difficult to resolve. One way we can try to determine the efficacy of the Federal Reserve's discount window lending is by examining how much less concentrated balances at the Federal Reserve were after discount window

loans were made. We can determine this difference by revisiting the HHI of account balances and comparing it with the concentration of account balances with the discount window loans removed. The discount window loans not only provided a high level of liquidity, but were distributed in such a way as to decrease the concentration of balances. We calculated that the discount window lending led to a distribution of balances across banks that was 8 percent less concentrated (as measured by the HHI) on September 11, and 9 percent less concentrated on the twelfth than the balances would have been if discount window lending had not taken place.

The redistribution of balances and the injection of new balances via the discount window are consistent with the interpretation that the loans allowed banks to carry out payments. By assuring banks that the discount window was available and by waiving daylight overdraft fees, the Federal Reserve's actions tended to eliminate the strategic elements of the timing and submission of payments.

We argue that the Federal Reserve System's actions were likely effective in facilitating payments during the period of technical disruption and in affecting banks' abilities to make and coordinate payments. It is important to note that the Federal Reserve System's injections of liquidity, and other actions urging banks to make payments, did not resolve the physical problems that banks had in communicating or accessing customer records. Instead, they were aimed at affecting the behavior of banks that were grappling with both the technical difficulties and their resulting liquidity problems.

The discount window loans not only provided a high level of liquidity, but were distributed in such a way as to decrease the concentration of balances.

The Discount Window, Open Market Operations, and the Coordinator of Last Resort

The Federal Reserve System injected liquidity, primarily through open market operations, in the period starting on September 13. On that day, the Federal Reserve advanced banks a much smaller amount in discount window loans (but still far more than average), and on September 14, its discount window lending was at a normal level. On those same days, the Federal Reserve injected a large amount of liquidity via open market operations.

The switch from primarily using the discount window to relying more on open market operations and the fed funds market occurred after some of the major technical disruptions were resolved, communications improved, and more normal patterns of coordination in payments activity had begun to be reestablished. Without a fully operating fed funds market, open market operations are unlikely to be effective (relative to the use of the discount window) in channeling funds to the institutions that need funds to send payments. Another factor that distinguishes the discount window as a means of injecting liquidity is that it is available to banks upon their request. Under this facility, Federal Reserve Banks may make credit available to depository institutions by providing advances secured by acceptable collateral or by discounting paper that meets the requirements of the Federal Reserve Act (see Board of Governors of the Federal Reserve System [2002] for more discussion of the discount window). In contrast, open market operations are made at the discretion of the central bank, and the liquidity that is injected must typically be redistributed to reach a particular bank's account.

Even when the fed funds infrastructure is fully intact, there are reasons to believe that the interbank market will perform poorly when payments are uncoordinated.[36] As we have seen, such situations lead to a generally increased demand for balances. Banks that are reluctant to pay one another are also likely to be reluctant to lend to one another. In these cases, injecting funds through open market operations may not be effective because the funds may not be circulated to the particular banks that most need liquidity. Discount window operations may be more effective in reestablishing coordination, but once coordination has largely been reestablished, open market operations are, in general, preferable because the operation of the fed funds market can be expected to distribute balances in an efficient way.

It is notable that much of the recent literature on the discount window, such as Schwartz (1992), Goodfriend and King (1988), and Furfine (2000), either overlook this motive for the use of the discount window or dismiss it in their evaluations as probably unnecessary (for an alternative view, see Martin [2002]). In all three articles, the authors assume that payments coordination is not an issue. Furthermore, the authors assume that the fed funds market, combined with open market operations, can provide and direct adequate liquidity to

the banking system. Neither of these assumptions is one that can be made for all feasible circumstances.

In general, no single private bank can appropriate all of the spillover public benefit of the coordination, so the incentive to reestablish coordination by borrowing at the target funds rate is likely to be too low relative to the social optimum. Therefore, the central bank is in a unique position to assist banks in recoordinating payments. That the Federal Reserve waived daylight overdraft fees and penalty fees on overnight overdrafts and assured banks that liquidity was available from the discount window is consistent with providing liquidity at rates that take into account the social benefits of payments coordination. In its role as "coordinator of last resort," the central bank might use either open market operations to inject funds (as long as the interbank market is functioning well) in sufficient amounts for the interbank rate to trade at a discount to the target rate, or it might use the discount window.

Discount window operations may be more effective in reestablishing coordination, but once coordination has largely been reestablished, open market operations are . . . preferable because the operation of the fed funds market can be expected to distribute balances in an efficient way.

In their September 17, 2001, statement in which they lowered the fed funds target rate and the discount rate, the Federal Open Market Committee (FOMC) and the Federal Reserve Board announced, "the Federal Reserve will continue to supply unusually large volumes of liquidity to the financial markets, as needed, until more normal market functioning is restored. As a consequence, the FOMC recognizes that the actual federal funds rate may be below its target on occasion in these unusual circumstances."[37]

Historically, the discount rate has been set below the target funds rate.[38] The important point from the coordination-of-payments perspective is that the Federal Reserve can capture some of the public benefit from reestablishing the coordination equilibrium.[39] For example, once coordination is reestablished, the number of intraday overdrafts declines, reducing the risk taken on by the Fed.

Payments System Infrastructure and Design

It is worthwhile to ask if central banks could make changes to the operation of payments systems that would prevent or attenuate problems that tend to disrupt the coordination of payments timing. An important change that market participants and financial system regulators are actively pursuing is to work to safeguard the physical infrastructure used in the payments system. The Federal Reserve, along with other financial system regulators, has pursued this approach in the recent "Draft Interagency White Paper on Sound Practices to Strengthen the Resilience of the U.S. Financial System." In that paper, the regulators identify sound practices to strengthen the U.S. financial system's resilience. These practices include, among others, a practice for firms that play significant roles in critical markets to "maintain sufficient out-of-region resources to meet recovery and resumption objectives."

In addition to these fundamental changes in the payments infrastructure, there could also be changes in the protocols for submission and settlement of payments. These changes might occur in ways that would either assist banks in maintaining or reestablishing coordination in the case of an interruption to the normal patterns of payments flows, or might reduce the reliance of banks on payments coordination for liquidity purposes altogether. For example, McAndrews and Rajan (2000) review a number of options to improve the coordination of payments in the normal course of events, such as offering zero-cost overdraft privileges during specified times every day to encourage banks to submit payments during those times. Such an approach might be useful in providing a more targeted focal point to reestablish coordination of payments, but it does not reduce the likelihood of a coordination failure.

Alternatively, the payments system could be designed to eliminate the strategic elements of timing and submission of payments altogether. Several options are possible. One option is to implement new "hybrid" payments systems—an example of which is the new CHIPS system in the United States—that seek to lessen the motive for timing payments to save on liquidity costs.[40] Pure netting systems accomplish this as well, but at the cost of delays in payments settlement until the designated time and also at the cost of creating intraday credit exposures that remain outstanding until settlement. It is generally agreed that deferred net settlement systems cannot accommodate the requirements of modern financial systems, which require intraday finality of payments, or the settlement of payments throughout the day.

The hybrid payments systems offer an alternative in that they allow for some subset of the payments to be settled during the day, while some other subset remains in a queue of payments awaiting settlement. Various possibilities for settling payments from the queue are possible, but they tend to share a key feature, which is that a bank is strictly better off submitting payments to the queue than delaying a submission until it receives a payment from another bank. Typically, the operator of the payments system can search the queue for offsetting payments. A bank is then better off submitting its payments to the queue because by doing so, the bank makes the settlement of the payment contingent on the receipt of an offsetting payment. Such a system, therefore, offers the possibility that a bank will submit payments to the queue early in the day, letting the operator of the queue management system release that bank's payments when messages of received payments arrive.

It is worth investigating whether such systems, if implemented, would operate well in distributing liquidity and in encouraging the submission of payments, even if one bank were unable to send out payments.[41, 42] These systems typically have been considered a means for banks to economize on liquidity usage (while achieving intraday settlement of payments) during normal operations. Given the magnitude of the physical disruptions experienced on September 11, 2001, it is now clear that an important consideration in the design of a payments system is how that system would perform under such extreme circumstances.

Conclusion

Over the years, the interdependency of payments has become more important and prominent in large U.S. dollar payments for two reasons. First, the turnover on Fedwire has been increasing for years, highlighting the importance of incoming funds as a source of liquidity for payments, and thus creating greater demand for coordination. Second, the level of the concentration of payments has been rising, increasing the ability of the relatively unconcentrated payments providers to coordinate their activities. As coordination has become more important for the normal functioning of the payments system, the breakdown of coordination has become more problematic in that the demands for liquidity balloon as the coordinating equilibrium fails.

In this article, we have shown how the Federal Reserve System acted to restore payments coordination through especially high levels of discount window and intraday lending on September 11 and 12. As communications improved, the Federal Reserve was able to meet more of the market's needs for liquidity through open market operations, allowing the fed funds market to distribute that liquidity. We have also argued that because of the social benefit of reestablishing the payments coordination equilibrium, the injection of funds into the U.S. payments system was important in overcoming the breakdown of the equilibrium in payments coordination. Going forward, it is likely that the discount window will continue to provide an important way for the Federal Reserve to direct liquidity to banks and to improve the coordination of payments in periods of severe disorder in the pattern of those payments.

Endnotes

1. Fedwire is a real-time gross settlement (RTGS) payments system. In such systems, payments are executed and finalized very shortly after they are communicated by the originating bank to the system operator and after the system operator communicates the payments details to the receiving bank.

2. See Bank for International Settlements (1997, p. 22) for a similar set of sources of funding. In this article, we refer to "liquidity" as the ability of banks—whose balance sheets often consist of loans and other assets that may be difficult to sell in the short run—to make outgoing payments immediately. As a result, liquidity is a more general concept than "funds in a central bank account." It includes the expected behavior of other participants in the payments and banking markets, and arises as part of an equilibrium of behavior of the market participants.

3. Service-related balances are primarily made up of required clearing balances, which banks establish to conduct payments operations via Fedwire when their reserve balances are too low to accommodate their payments needs. See Stevens (1993) for a description of required clearing balances.

4. As these values trend upward with growth in economic and financial activity, August 2001 is used as an example of recent activity prior to the period of interest. See Board of Governors of the Federal Reserve System (2001c, p. A5, Table 1.11, lines 22 and 25).

5. An exception to this rule should be noted: required clearing balances earn credits that can be used to offset fees charged by the Federal Reserve System for its priced services. See Stevens (1993) for a more complete discussion.

6. As explained by Furfine (1999), such a large amount of lending is possible, given the much lower balances available to lend, by banks acting as both borrowers and lenders on the same day in the fed funds market on a regular basis.

7. There is a third possibility, previously mentioned: the bank can run an "overnight overdraft" for which it pays a 4-percentage-point penalty above the effective federal funds rate for that day.

8. See Coleman (2002, p. A6, Table 1.12, line 8).

9. See Board of Governors of the Federal Reserve System (2002) for the recent proposal regarding the discount window.

10. See Coleman (2002) for an extensive discussion of the Federal Reserve's intraday credit policies and operations. Not all banks can borrow intraday from the Fed, and the borrowing of any particular bank is subject to a limit on the amounts that can be borrowed. The Fed guarantees payment of funds transfers across Fedwire, so if a bank has a negative balance in its account and is within its borrowing limit, the Fed transfers the funds to the receiving bank, effectively lending the "daylight overdraft" to the bank that originated the payment.

11. See Coleman (2002, p. 82, Table 8).

12. The turnover ratio has increased in recent decades, both in the United States and abroad. See, for example, Bank for International Settlements (1997).

13. See, for example, Angelini (1998), Bech and Garratt (forthcoming), and Kobayakawa (1997).

14. See Bulow et al. (1985) for a general description of strategic complementarity. Strategic complementarity is indicative of situations in which agents' payoffs increase with the degree of coordination with other players.

15. Of course, in a deferred (or designated-time) net settlement payments system, offsetting payments made at different times are cumulated and offset at the time designated for settlement. Another way banks can use incoming funds for payments in an RTGS system is by using "throughput guidelines," which require banks to submit certain percentages of the payments the banks make for the whole day by specified times during the day. Such throughput requirements can regulate the rate of turnover of account balances and are in use in the United Kingdom's RTGS system, the Clearing House Automated Payments System (CHAPS). Throughput requirements are agreed upon by all system members and therefore might be interpreted as a cooperative means of coordinating payments flows.

16. McAndrews and Rajan (2000) present evidence consistent with this hypothesis for Fedwire payments activity.

17. On this point, see Angelini (2000), Kahn and Roberds (1998), and Kahn, McAndrews, and Roberds (forthcoming).

18. Once again, an equilibrium of payments submissions may result in a delayed (relative to the case in which there is no risk of default by the counterparty), but simultaneous, submission of payments.

19. The Herfindahl-Hirschman index is the sum of the squares of the market shares of the deposits of the banks. It can vary from 0 to 10,000, with 10,000 indicating that a single bank holds all of the deposits. For example, a level of 5,000 corresponds to a symmetric duopoly in deposit holdings and a level of 3,333 corresponds to a symmetric triopoly.

20. Specifically, the netting ratio over the period from t_1 to t_2 is defined as:

$$N = \left[\frac{\sum_{s=t_1}^{t_2} \sum_i \sum_j P_{ij}^s}{\frac{1}{2}\left\{ \sum_{s=t_1}^{t_2} \sum_i \left| \sum_j P_{ij}^s - \sum_j P_{ij}^s \right| \right\}} \right],$$

where P_{ij}^s represents the value of the payment from bank i to bank j in time period t.

21. The example captures many features of the situation under review. For example, although a bank can generally make a payment if its balance at the Federal Reserve is zero (by borrowing from the Federal Reserve via a daylight overdraft), this is a costly way of making a payment if the bank is expecting other payments to arrive. It is also risky to make a payment by borrowing while awaiting the arrival of other payments because if the other payments were not to arrive, then the bank could incur the substantial penalty of an overnight overdraft. As a result, the failure of an expected payment to arrive can cause another bank to delay sending its payments, which are expected by yet other banks, in turn causing them to delay their payments, and so on. This multiplier effect of the original delay can cause all other banks in the system to demand liquidity.

22. The *Financial Times*, on September 21, 2001, reported that the Bank of New York was the hardest-hit bank because its two principal locations are in Lower Manhattan.

23. Fedwire's operations were uninterrupted on September 11 and thereafter.

24. One of these estimates is gathered by the Markets Group of the Federal Reserve Bank of New York, in its daily survey of fed funds brokers. The other estimate was made by the authors, using a method suggested by Furfine (1999). His method consists of systematically culling all large interbank payments from the records of all Fedwire funds transfers that satisfy certain criteria. Among those criteria are that the selected payment's values must be at least $1 million and be made in increments of $100,000. In addition, each selected payment must match a payment made the following banking day in the reverse direction, and whose value exceeds the selected payment by an amount that would closely correspond to the interest that would be expected given the range of reported fed funds interest rates for that day.

25. The *benchmark* data set includes one day for each ten-day maintenance period from May 1999 to August 2001. In addition, we include the weeks of September 4, 2001, to September 10, 2001, and September 24, 2001, to November 28, 2001, for 107 days in all. The remaining data from September 11, 2001, to September 21, 2001, are used for a detailed analysis of activity on Fedwire on the day of the attacks and the eight business days thereafter.

26. If a large proportion of the balances in the banking system concentrate in one bank's account, then other banks will face, all else being equal, higher costs of making payments, or alternatively may face liquidity constraints on their borrowing, which could preclude their submission of further payments. As shown in Exhibit 2, we use the Herfindahl-Hirschman index to measure concentration of balances in Federal Reserve accounts.

27. Fedwire hours of operation were extended during the first week following the attacks; Fedwire's usual closing is 6:30 p.m.

28. Coleman (2002) reports that daylight overdrafts increased from their August 2001 average of $32.8 billion to $45 billion on September 11, $36 billion on the twelfth, $41 billion on the thirteenth, and $54 billion on the fourteenth.

29. Other formulations, more directly comparable with the exact form of the best-reply function of Bech and Garratt (forthcoming), were tested, such as estimating a bank's outgoing payments in response to all other banks' outgoing payments. Similar qualitative results were obtained from those formulations.

30. We estimated this model using receipts both from other banks' funds transfers, as well as the settlement of sales of a bank's government and agency securities, which are also settled via Fedwire. Including these receipts in the estimation had very little effect on the estimates on receipts from other banks' funds transfers and the other variables in the model.

31. Few banks make payments in the early morning hours. Banks may make "settlement" payments to other banks after 6:00 p.m., but not payments on behalf of customers. Dummy variables for the other days of the reserve maintenance periods were included in other estimations, but were not significant.

32. It should be noted that the results are robust with respect to the exact times over which one measures outgoing payments and receipts.

33. The Bank of New York (2001) issued a press release on September 14 announcing that "virtually all of its systems are up and running."

34. See Board of Governors of the Federal Reserve System (2001a).

35. The Federal Reserve System announced the waiving of daylight overdraft fees via a Fedwire broadcast message on Friday, September 14.

36. In Europe, where no attacks or technological problems occurred, there was, nonetheless, significantly higher demand for liquidity on September 12, indicated by a spike in interest rates on overnight interbank loans. The European Central Bank supplied an extraordinary amount of liquidity on September 12 and 13. See Bindseil et al. (2002) for a more complete discussion.

37. See Board of Governors of the Federal Reserve System (2001b).

38. See Board of Governors of the Federal Reserve System (2002) for the Board's May 17 proposal to change discount window lending policies to set the discount rate higher than the target fed funds rate.

39. The public benefit to higher coordination of payments includes the lower cost of payments, as banks can hold lower overnight balances to complete expected payments, which reduces the "inflation tax" to which the public is subject for holding currency or noninterest-bearing balances at the central bank. In addition, banks are subject to either less risk or a shorter duration of risk exposure from their counterparties.

40. McAndrews and Trundle (2001) and Bech and Soramäki (2001) describe various hybrid systems.

41. One area of interesting research would be to examine the performance of CHIPS relative to Fedwire following the terrorist attacks.

42. In the United States, for example, many banks can choose between Fedwire and CHIPS to make payments. A simulation analysis of the effects of infrastructure disruptions in a payments system is presented in Bech and Soramäki (2002).

References

Angelini, Paolo. 1998. "An Analysis of Competitive Externalities in Gross Settlement Systems." Journal of Banking and Finance 22, no. 1 (January): 1-18.

———. 2000. "Are Banks Risk Averse? Intraday Timing of Operations in the Interbank Market." Journal of Money, Credit, and Banking 32, no. 1 (February): 54-73.

Bank for International Settlements. 1997. "Real-Time Gross Settlement Systems." Committee on Payment and Settlement Systems Publications no. 22, March.

Bank of New York. 2001. "Virtually All Bank of New York Systems Are Up and Running." Press release, September 14.

Bech, Morten, and Rod Garratt. Forthcoming. "The Intraday Liquidity Management Game." Journal of Economic Theory.

Bech, Morten, and Kimmo Soramäki. 2001. "Gridlock Resolution in Interbank Payment Systems." Bank of Finland Discussion Paper no. 9/2001.

———. 2002. "Liquidity, Gridlocks, and Bank Failures in Large Value Payment Systems." E-money and Payment Systems Review, Central Bank Publications, London.

Bindseil, Ulrich, Benedict Weller, and Flemming Wuertz. 2002. "Central Bank and Commercial Banks' Liquidity Management—What Is the Relationship?" May. European Central Bank.

Board of Governors of the Federal Reserve System. 2001a. "Federal Reserve Is Open and Operating." Press release, September 11.

———. 2001b. "FOMC Statement; and Board Approval of Discount Rate Requests of the Federal Reserve Banks of Richmond, Chicago, Minneapolis, Dallas, and San Francisco, Effective Immediately." Press release, September 17.

———. 2001c. Federal Reserve Bulletin 87, no. 11 (November): 719-28.

———. 2002. "Proposal to Revise Federal Reserve Discount Window Programs." Press release, May 17.

Board of Governors of the Federal Reserve System, Office of the Comptroller of the Currency, and Securities and Exchange Commission. 2002. "Draft Interagency White Paper on Sound Practices to Strengthen the Resilience of the U.S. Financial System." Docket no. R-1128, August 30.

Bulow, Jeremy, John Geanakoplos, and Paul Klemperer. 1985. "Multimarket Oligopoly: Strategic Substitutes and Complements." Journal of Political Economy 93, no. 3 (June): 488-511.

Coleman, Stacy Panigay. 2002. "The Evolution of the Federal Reserve's Intraday Credit Policies." Federal Reserve Bulletin 88, no. 2 (February): 67-84.

Diamond, Douglas W., and Philip H. Dybvig. 1983. "Bank Runs, Deposit Insurance, and Liquidity." Journal of Political Economy 91, no. 3 (June): 401-19.

Federal Deposit Insurance Corporation. 1997. History of the Eighties—Lessons for the Future. Volume 1: An Examination of the Banking Crises of the 1980s and Early 1990s. Washington, D.C.

Federal Reserve Bank of New York. 2002. Annual Report.

Furfine, Craig. 1999. "The Microstructure of the Federal Funds Market." Financial Markets, Institutions, and Instruments 8, no. 5 (November): 24-44.

———. 2000. "Empirical Evidence on the Need for a Lender of Last Resort." April. Bank for International Settlements.

Goodfriend, Marvin, and Robert King. 1988. "Financial Deregulation, Monetary Policy, and Central Banking." Federal Reserve Bank of Richmond Economic Review 74, no. 3 (May/June): 3-22.

Group of Ten. 2001. "Report on Consolidation in the Financial Sector." Bank for International Settlements.

Kahn, Charles, James McAndrews, and William Roberds. Forthcoming. "Settlement Risk under Gross and Net Settlement." Journal of Money, Credit, and Banking.

References (Continued)

Kahn, Charles, and William Roberds. 1998. "Payment System Settlement and Bank Incentives." Review of Financial Studies 11, no. 4: 845-70.

———. 1999. "The Design of Wholesale Payments Networks: The Importance of Incentives." Federal Reserve Bank of Atlanta Economic Review 84, no. 3: 30-9.

Kobayakawa, Shuji. 1997. "The Comparative Analysis of Settlement Systems." Centre for Economic Policy Research Discussion Paper no. 1667, July.

Martin, Antoine. 2002. "Reconciling Bagehot with the Fed's Response to September 11." Unpublished paper, Federal Reserve Bank of Kansas City, October.

McAndrews, James, and Samira Rajan. 2000. "The Timing and Funding of Fedwire Funds Transfers." Federal Reserve Bank of New York Economic Policy Review 6, no. 2 (July): 17-31.

McAndrews, James, and John Trundle. 2001. "New Payment System Design: Causes and Consequences." Bank of England Financial Stability Review 11, December: 127-36.

Schwartz, Anna. 1992. "The Misuse of the Fed's Discount Window." Federal Reserve Bank of St. Louis Review 74, no. 5 (September/October): 58-69.

Stevens, Ed. 1993. "Required Clearing Balances." Federal Reserve Bank of Cleveland Economic Review, December: 2-14.

Volcker, Paul A. 1986. "Operational Problems Experienced by the Bank of New York Resulting from a Failure of Computer System Software." Statement to the U.S. House Subcommittee on Domestic Monetary Policy of the Committee on Banking, Finance, and Urban Affairs, December 12, 1985. Reprinted in Federal Reserve Bulletin 72, no. 2 (February): 115-6.

Jason Bram, Andrew Haughwout, and James Orr

Has September 11 Affected New York City's Growth Potential?

- Over the past three decades, New York City's economy has shown many signs of strength, including increased worker earnings and rising land prices.
- Although the attack caused a sharp temporary disruption in the economy, an advantageous industry mix—one weighted toward high-paying, rapidly expanding industries—is likely to keep the city well positioned for growth over the medium term.
- Still, if the city is to translate favorable prospects into actual growth, it must maintain an environment that is attractive to firms and workers. Thus, it must rebuild its damaged infrastructure and close a sizable budget deficit without letting services deteriorate or taxes rise too high.
- Preliminary evidence suggests that the demand for New York City property remains robust. Continued strength in land prices will be an important measure of the city's growth.

The terrorist attack on the World Trade Center claimed close to 3,000 lives and caused billions of dollars in property damage. In the aftermath, New York City's economy contracted briefly but sharply: many businesses were forced to shut down, mostly temporarily, and tens of thousands of workers were either dislocated for a short time or lost their jobs (Bram, Orr, and Rapaport 2002). The economic implications for the city, however, clearly extend beyond the first few months following the attack.

In this article, we analyze the effects of September 11 on the longer run prospects for the New York City economy. We find that, on the one hand, several downside risks to the city's growth outlook have arisen. In a worst-case scenario, the concentration of the attack on Lower Manhattan has raised the possibility that financial firms might relocate outside the city, which could generate a cumulative downward spiral of job and income growth. On the other hand, the city's industrial structure and its quality-of-life amenities—namely, an industrial mix weighted toward high-growth sectors and an environment desirable to workers and firms—suggest favorable economic prospects. However, whether or not these prospects translate into actual growth going forward depends to a large degree on the city's policy response to the economic pressures arising from the attack. The key elements of this response will be the ability to avoid budgetary decisions that

Jason Bram is an economist, Andrew Haughwout a senior economist, and James Orr a research officer at the Federal Reserve Bank of New York.

The authors thank two anonymous referees, as well as Nathan Barczi for valuable research assistance. The views expressed are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System.

reduce the long-run attractiveness of a New York City location and, through rebuilding, to maintain the productivity of the city's capital stock.

We begin by describing recent earnings and land price trends in the city. To help interpret these data, we apply a model that emphasizes the importance of local property markets as an indicator of trends in a mature urban economy. The New York experience is then discussed in relation to the model, and the city's economic strengths and weaknesses both before and after the attack are evaluated. We also identify several challenges that local policymakers will face if the city is to recover fully. Finally, our study presents evidence on land prices after September 11 indicating that a strong demand for New York City property still exists.

Growth Trends in New York City

Three important patterns characterize the trends in the New York City economy over the past three decades: steady but cyclical employment, rising real earnings, and appreciating land prices.

Economic growth at the national level is usually measured as the average annual rate of expansion of real (inflation-adjusted) GDP over some period. At the city level, however, an official output measure is not available and thus growth is often measured as the trend rate of growth in jobs and/or real income. Although in many localities these measures move together, the measures in New York City present different perspectives on the process of economic growth. Average total employment in the city essentially has been unchanged for three decades (Chart 1). Even with the rapid expansion of jobs in the mid-to-late 1990s, only in 1999 did total employment exceed its previous cyclical peak, in 1989. Indeed, the all-time peak level of total employment was reached in 1969.

Three important patterns characterize the trends in the New York City economy over the past three decades: steady but cyclical employment, rising real earnings, and appreciating land prices.

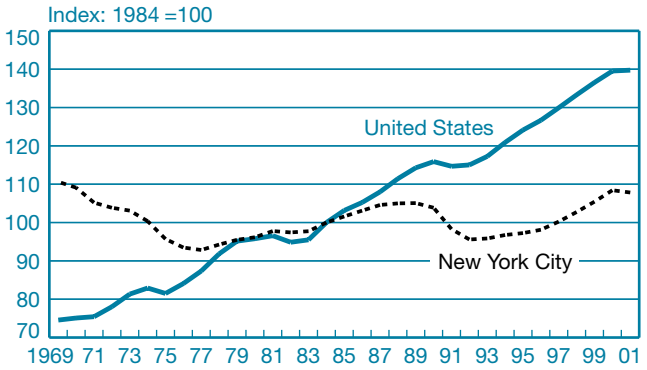
In contrast, real earnings of city workers have trended upward since 1980 at an average annual rate of about 3.5 percent.

New York's pattern of steady employment and rising real earnings is attributable to occupational and industrial restructuring. Average earnings in the city have risen because of accelerating productivity growth in existing jobs and a shift toward higher paying jobs. During the 1980s, the expansion of jobs in the high-paying FIRE (finance, insurance, and real estate) sector helped raise average real earnings citywide. Many of these job gains were temporarily reversed in the recession of the early 1990s. However, the city's recovery since the mid-1990s, in conjunction with the job expansion in some relatively high-paying service sectors, again boosted real earnings.

The city increased its earnings relative to the rest of the country by enhancing the productivity of existing employment and capturing a rising share of high-productivity jobs.

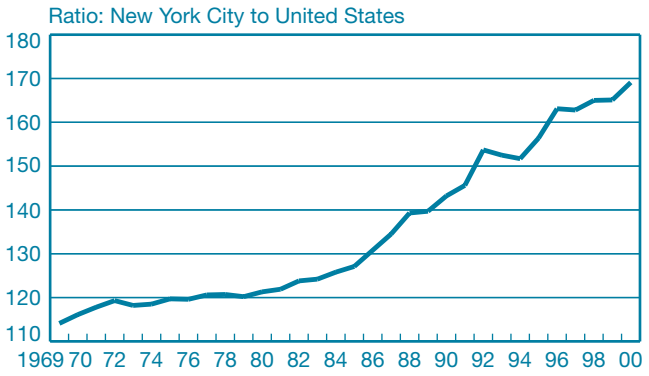
Over this same period, earnings per job increased nationwide, but New York's earnings per job rose even more rapidly (Chart 2). Thus, the city increased its earnings relative to the rest of the country by enhancing the productivity of existing employment and capturing a rising share of high-productivity jobs.

Chart 1
Total Employment in New York City and the United States



Source: U.S. Department of Labor, Bureau of Labor Statistics.

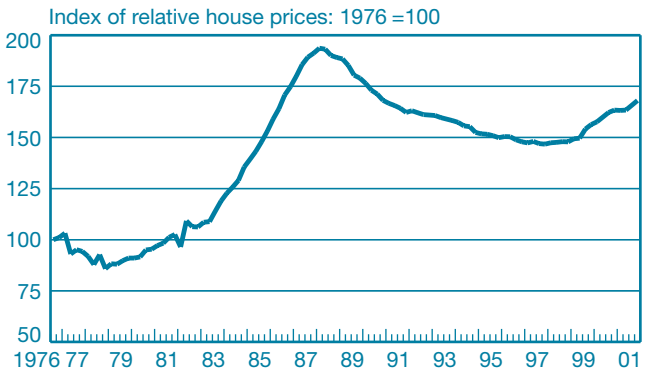
Chart 2
Relative Earnings per Worker in New York City



Sources: U.S. Department of Commerce, Bureau of Economic Analysis (earnings); U.S. Department of Labor, Bureau of Labor Statistics (total employment).

Note: Earnings are nominal.

Chart 3
New York City Area House Prices Relative to U.S. Average



Sources: Office of Federal Housing Enterprise Oversight; Federal Reserve Bank of New York calculations.

Notes: The index is based on the ratio of the repeat-sales price measure for existing single-family homes in the New York City metro area to that of the United States overall; the index is designed to control for changes in the mix of homes sold. Data for the city itself are not available, so we present the index for the New York City primary metropolitan statistical area (PMSA), which consists of the city's five boroughs (the Bronx, Brooklyn, Manhattan, Queens, and Staten Island) and Westchester, Putnam, and Rockland counties in New York State. Approximately 80 percent of the population of the New York PMSA lives within the city boundaries.

Furthermore, the price of housing in New York has also risen relative to the nation since 1976 (Chart 3). Here, we use the repeat-sales price index, which controls for the quality of the structure, as a measure of price change. Constant-quality housing price changes provide a relatively clean measure of the attractiveness of the bundle of local traits available to residents of these homes. These data indicate that the price of residing in New York has climbed relative to the rest of the nation—with a particularly sharp rise in the second half of the 1990s.

Modeling Growth in a Mature City

To analyze the effects of the terrorist attack on the New York City economy, we adopt the model of urban economies developed by Roback (1982) and refined in Blomquist, Berger, and Hoehn (1988), Gyourko and Tracy (1991), and Haughwout (2002). In this framework, metropolitan areas are viewed as small, open economies to which labor and capital are elastically supplied (see the box for more details on the model). Since each city is just one of many places where firms and households may choose to locate, it must offer competitive levels of profit for firms and utility for households. The value of a city to firms and households determines their "bids" in the city's local land and labor markets. Thus, if a location has fixed characteristics that are extremely productive, firms will offer high bids for sites and high wages to attract workers there. Locations may vary in their attractiveness to firms for a variety of reasons, ranging from access to valuable existing infrastructure stocks to proximity to markets or sources of raw materials.

Given this set of productive amenities, firms will relocate if wages and/or land prices are too high for them to make equilibrium profits. If costs are so low that incumbent firms can make excess profits, new firms will enter, bidding up local prices. A similar logic applies to households. In a mixed economy that provides sites for households as well as firms, local land and labor prices are determined by the satisfaction of firm and household equilibrium conditions. The local price equilibrium is thus attained when no firm or household wishes to relocate.

When a city's productive or residential environment changes over time, such changes will be reflected in local prices. Improvement in a city's appeal to firms and/or households will lead to a rise in relative land prices over time.[1] When a city's attractiveness is at its highest to firms (for instance, if taxes paid by firms fall but nothing else changes), wages will tend to rise.

A Model of Local Prices in Mature Urban Economies

Free mobility means that firms and households must pay, albeit indirectly, for attractive local characteristics. Firms and households pay for local amenities through higher land prices and wages. Thus, when a place offers household amenities like an attractive climate or unique consumption opportunities, households will be willing to pay to locate near these amenities, bidding more for land there. Yet, since wages represent an income source to households, they will also be willing to accept *lower* wages to locate in places with features they value. If a place is a "bargain" to households in the sense that its land prices are low and its wages are high relative to its attractive amenities, households will move in, bidding land prices up and wages down. Firms, however, bid more for attractive sites in both land and labor markets.

Exhibit 1 presents equilibrium land and labor bids by a competitive firm (labeled $\Pi = \Pi_0$) and a representative household (labeled $V = V_0$) for a place with a given set of production and consumption amenities (A_0). Note that the firm curve (an isoprofit curve in the price space) is downward-sloping, indicating that as wages in a jurisdiction rise, land rents must fall if firms are to remain profitable. The household indifference curve in the price space is labeled *V*, and it slopes upward: higher wages are associated with higher land prices. With higher incomes, households must pay higher land rents or they will earn excess utility—leading to new households bidding in local land and labor markets. The point where both the firm and household equilibrium curves intersect is the local price equilibrium (W_0, R_0).

Exhibit 1
Wage and Land Rent Equilibrium

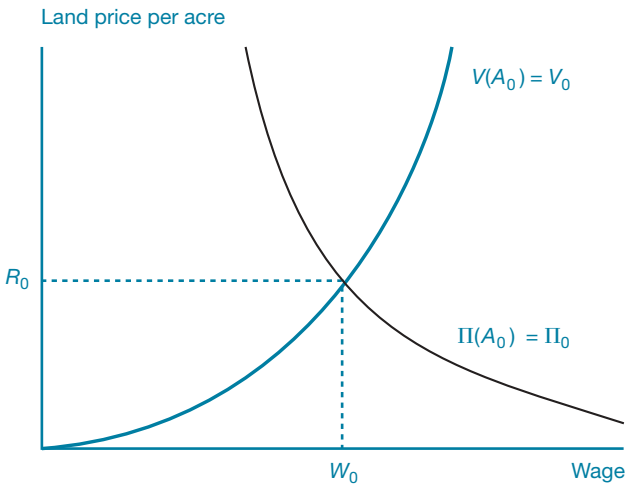


Exhibit 2
Wage and Land Rent Equilibrium with Fewer Amenities

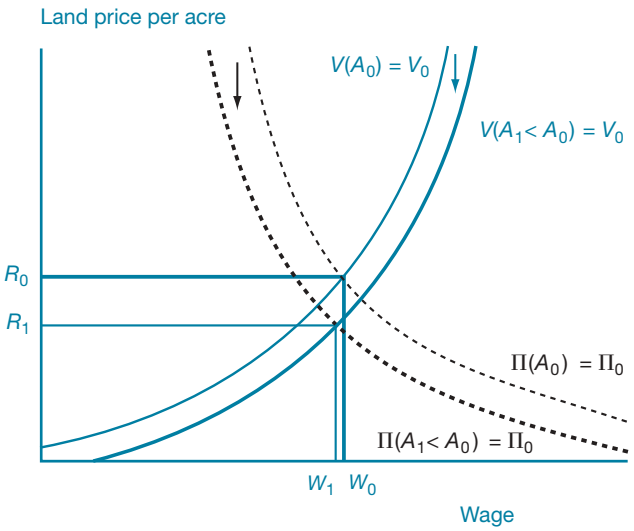


Exhibit 2 depicts the effect of a change in local amenities (to $A_1 < A_0$) on this equilibrium. When a city becomes less attractive to firms, the $\Pi = \Pi_0$ curve shifts downward. Changes that reduce the attractiveness of a location to households will shift the $V = V_0$ function downward. For given wage levels, firms and households are willing to pay less for land in unproductive places. Note that a reduction in attractiveness has an unambiguously negative effect on land prices, but that the effect on wages will depend on which curve (household *V* or firm Π) shifts more. When firms are disproportionately hurt by a change, the fall in equilibrium land rents will be accompanied by a *reduction* in wages. A relatively large impact on households would lead to *increases* in wages, as households demand higher compensation to induce them to reside and work in an unattractive, low-amenity place.[a]

[a]Haughwout (2002) formally derives these comparative statics and provides an example of the use of wage effects in the determination of the incidence of policy benefits.

When households receive the lion's share of the benefits of an amenity change (perhaps a costless reduction in the rate of violent crimes), relative wages would be expected to fall.

Wage and land price data for New York City indicate that the trends prior to September 11 were favorable, as both wages and relative land prices had been increasing over long and short horizons (Charts 2 and 3). Land price increases are a sign of growing attractiveness to firms and/or households, while wage increases indicate that businesses increasingly value a New York location relative to households. Interpretation of the wage increases observed in the city, however, is also confounded by the change in the population's characteristics over time, as we shall discuss.

Prices of housing and land, both of which are durable assets, reflect not just current conditions, but also market participant views of a location's future. The demand for a property is determined by the present value of the stream of net benefits expected to accrue to the property over its lifetime. When an individual purchases a piece of real property, he or she must evaluate future conditions in the area where the property is located. If conditions are expected to deteriorate next year, or in five years, then purchase prices this year will be reduced, as the stream of returns into the future falls. Relative property prices in any period thus indicate, in part, expected future conditions in that location. Note that this predictive power of property prices does not extend to either property rentals or wages—both of which reflect current conditions or, more precisely, conditions expected to pertain to the duration of the contract.

Our discussion provides some insight into the sources of the apparent land market strength in New York City. Among the factors explored in some detail in the academic literature are changes in local fiscal policies, such as a fall in local taxes without compensating service decreases or an increase in local infrastructure provision. On the household side, safety from crime, the quality of the public educational systems, and taxes are all considered important factors determining local quality of life and thus local land and labor prices (see Blomquist, Berger, and Hoehn [1988] and Gyourko and Tracy [1991]). Measured by several of these factors, conditions in New York have improved over the past thirty years, with especially sharp improvements occurring in the past decade.

Applying the Model to New York City

Here, we extend the model presented above to account for the various types of firms and households that are locating in a modern economy.

The model dramatically simplifies the structure of mature real-world economies, wherein many kinds of firms and households coexist, particularly in urban areas. As a practical matter, the way that firms and households sort themselves over space is relevant as an indicator of both evaluations of locations and the prospects of particular areas. Thus, the fact that New York's relative (constant-quality) home values have increased significantly versus those of the nation indicates that the city has become more attractive to households and/or firms over the past few decades.[2]

Still, the benefits of a New York City location are clearly not the same to all actors in the economy. New York's industrial and demographic structures differ from those of the nation because some groups are willing to pay more than others for the city's particular amenities. Most obviously, New York's long domination of financial services employment provides significant incentives for firms in that industry to locate there.[3] But this same feature is less attractive to producers in, say, the automotive industry. Examples of the city's critical household amenities include relatively easy access to high-skill service employment opportunities and a diverse set of consumption opportunities such as theaters, museums, and other cultural offerings. However, the city's public schools, by most criteria, fall short of their suburban counterparts on several crucial dimensions.[4]

The fact that New York's appeal as measured by its relative land prices is near an all-time high seems to be contradicted by population and employment figures that have fallen relative to the nation. However, New York's situation is more complex than simple figures demonstrate: the city has substituted high-paying, high-productivity jobs in a few industries for low-skill, low-paying jobs in others. On the household side, corresponding changes in the composition of the local population have also tended to support a strong housing market, as we shall explain.

Clearly, the city's future depends on the growth prospects of the types of firms and households that have revealed a preference for New York. If, for example, the nation's financial services industry were expected to decline significantly, the city's concentration of these jobs would become a liability rather than an asset. Likewise, if concentrations of high-skill individuals or immigrants were detrimental to employment or population growth, the city's prospects would be dimmed.

New York's future—and the effect of events such as September 11—will therefore depend on the answers to two important questions:

- Can New York continue to provide the amenities valued most by those industries vital to its economy?

- Will the industries that benefit most from these amenities continue to thrive at the national level?

Our analysis suggests that the answers point to favorable prospects for New York City, although the future will not be without policy challenges. Our evidence indicates that New York is increasingly specialized in the production of skill-intensive services and in the provision of residences for high-skill workers and immigrants. These factors have supported rising land prices and wages in the city over the past three decades.[5]

The City's Growth Fundamentals

According to the model we employ, the desirability of a city springs from its productive and residential amenities. We begin our discussion of New York City's growth fundamentals by highlighting two key improvements over the past two decades that have made the city attractive to firms and workers: an improved fiscal position and a reduced crime rate. These improvements reflect both explicit policy choices and a healthier local economy. Next, we examine the associated changes in the city's industry and population mixes by reviewing the city's industrial structure and considering the medium-term prospects for these industries. Trends in the financial services sector are emphasized, because its fortunes exert a strong influence on the city's overall performance. We find that the local economy tends to have a relatively high concentration of industries expected to lead the nation in growth over the next decade. Finally, by profiling the changes in the local labor force, we find a more advanced educational profile in the city than in the nation as a whole—an advantage that has increased over the past ten years.

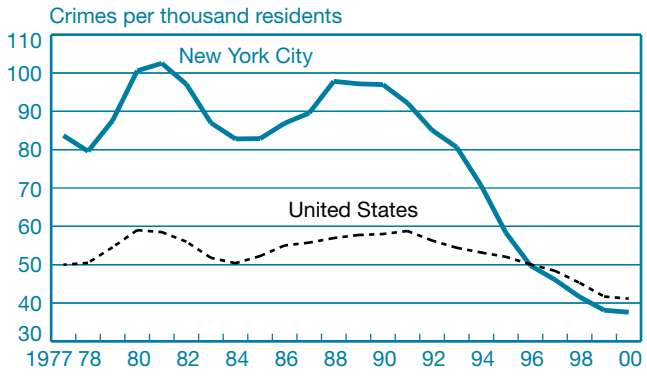
New York's Productive and Residential Environments

Both firms and households have benefited from the dramatic improvements in the city's financial condition since 1975, when New York experienced its most severe fiscal crisis since the Great Depression. The city has substantially reduced its reliance on debt finance, lowered property taxes, and enhanced its financial reporting since then (Haughwout 1997). Although the overall tax burden remains high, the share of the local tax dollar used to service short-term debt has been sharply reduced. Improvements in the city's public transportation system, particularly during the 1990s, also have presumably benefited firms and households.

Public safety is another key aspect of a location's attractiveness. In that regard, New York's crime rate tells a compelling story (Chart 4). The rate, which actually rose during the economic boom of the 1980s, began a steady and steep decline in 1991 (amidst one of the city's worst recessions) that continued through the end of the decade. Although the 1990s saw a downward trend in crime nationwide and the city's improving economy clearly helped to lower the crime rate, New York's decline in crime was even more pronounced. Moreover, neighborhood-specific data from New York City suggest that the sharpest declines in crime over this period tended to occur in the poorest areas and highest crime areas. The U.S. Census Bureau's periodically conducted New York City Housing and Vacancy Survey (HVS)—which includes a resident-reported measure of "neighborhood quality"—indicates that city dwellers perceived significant improvements in crime in their own areas, with the sharpest improvements noted in the poorest neighborhoods.

Although these improvements in the business and residential environments have helped to enhance relative land and labor prices, New York's portfolio of amenities is not without certain liabilities. Perhaps the most important of these are the city's relatively high taxes on residents and businesses, and public schools that fall short of suburban competitors. Nonetheless, the evidence suggests that these liabilities have been outweighed by improvements in recent decades.

Chart 4
Annual Crime Rate of New York City and the United States



Sources: Federal Bureau of Investigation; New York State Division of Criminal Justice Services.

Industry Mix

A city's industry mix indicates which sectors benefit most from its particular package of amenities and helps determine medium-term growth prospects. New York has long had a significant concentration of employment in the financial services industry. Firms in this sector appear to strongly value proximity to one another, implying that New York's main advantage in this sector has promoted its growth over the past several decades.[6]

In addition, areas with a high concentration of growth industries tend to benefit from trends in the broader economy. Accordingly, we examine New York City's current industry mix, which has resulted in part from the aforementioned changes in the city's environment. We find that the current mix is strongly weighted toward growth sectors. By comparing the performance of these industries in the city with their national counterparts, we find that most local industries have tended to lag moderately in terms of income growth but fairly substantially in terms of job growth. This finding is consistent with the model of New York City as a mature economy, with less potential for expansion than most parts of the nation—if for no other reason than land and space constraints.

In recent years, industries with a high concentration in New York City have tended to register relatively brisk job growth at the national level and are projected to continue doing so over the next decade.

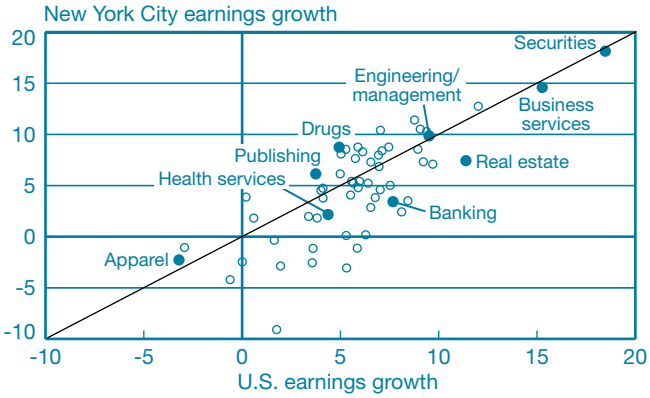
In recent years, industries with a high concentration in New York City have tended to register relatively brisk job growth at the national level and are projected to continue doing so over the next decade. In particular, the securities, business services, motion picture, legal, social, educational, and management services industries have all registered solid job and earnings growth at the national level, and their attraction to New York City has contributed to strong overall gains for the city. Conversely, the industries with the weakest trends in employment and earnings—almost all of which are in the manufacturing sector—are, for the most part, currently underrepresented in New York City. One notable exception is apparel manufacturing, which, though far less important than it was a few decades ago, is still one of the city's key industries.

Here, we illustrate the overall net effects of these industry dynamics by separating out the marginal effects of *local factors* and *industry mix* on overall income and job growth. Specifically, we decompose the differential between local and national growth into two components. *Local-factor effects* represent the performance of local industries compared with their national counterparts; these effects represent the portion of the overall growth differential not explained by local industry mix. *Industry-mix effects* represent what the differential would be if local job growth matched national job growth exactly within each industry; these effects represent the extent to which growth is enhanced or diminished by virtue of the local area's particular industry mix.

We conduct this exercise by first looking at changes in earnings over the 1995-2000 period to gauge the dynamics of the city's recent economic boom. Chart 5 presents local relative to national earnings growth for most two-digit Standard Industrial Classification industries, with selected key industries highlighted. As we see, some industries (those above the line) grew faster locally than they did nationally, while others (those below the line) grew more slowly. On balance, though, New York City's industries lagged their national counterparts in growth by an estimated 0.6 percentage point. However, because some of the nation's fastest-growing industries were

Chart 5

Earnings Growth by Industry: United States versus New York City, 1995-2000



Sources: New York State Department of Labor; U.S. Department of Labor, Bureau of Labor Statistics.

Notes: Industries above the diagonal line registered faster earnings growth in New York City than they did nationally; those below the line experienced below-average growth in New York City. The dark circles represent selected two-digit Standard Industrial Classification industries that are important to the New York City economy.

overrepresented locally, New York City benefited disproportionately from its industry mix, as aggregate earnings growth was boosted by an estimated 2.3 percentage points (see table).

In terms of employment growth between 1995 and 2000, these effects were somewhat less pronounced, but they generally worked in the same direction. Here, positive industry-mix effects fully offset negative local-factor effects, leaving overall local job growth on a par with that of the nation. Over this five-year period, the local-factor effect in earnings per job was very near zero. That is, all of the city's relative growth in earnings per job was attributable to industry mix.

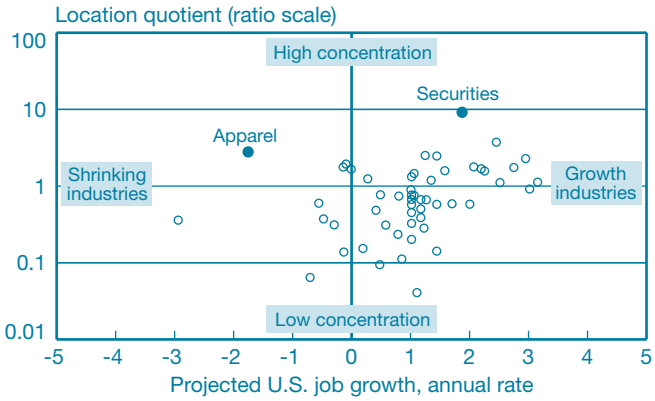
To gauge the likely impact of industry mix going forward, we performed a similar analysis using Bureau of Labor Statistics *projected* job growth rates by industry. The projected rates are for 2000-2010; comparable rates would presumably apply for 2002-12. Although we cannot predict local-factor effects, based on this information, industry-mix effects should continue to contribute positively, albeit modestly, to local job growth, as demonstrated in the table. As Chart 6 illustrates, industries with above-average expected job growth tend to be more concentrated in the city than those with subpar growth or job losses. Still, to the extent that local industries continue to lag their national counterparts in job growth, it is unclear whether local job growth would keep pace with the national average.

Because New York City has been moving toward increasingly high-value-added (and high-salary) industries and businesses over time, a comparison of local employment trends with corresponding national trends tends to understate the city's true relative performance. In fact, a study by McCarthy and Steindel (1997) points to the metropolitan region's persistently strong income growth as an indication of healthy demand for its goods and services. The authors argue that if income is holding up fairly well, then spending on regionally produced goods and services is high and business in the region may be more robust than the employment data suggest. As depicted in Chart 2, average earnings per worker have consistently grown more strongly in New York City than nationwide; between 1995 and 2000, this growth was due entirely to *industry mix*. If these long-standing trends persist, New York City should continue to fare better in terms of income growth than in job growth.

Looking forward, we note that to the extent that local-factor effects on earnings and employment remain neutral or small—a condition that depends on New York City remaining an attractive site for business locations—the city may be expected to continue its long-standing pattern of rising earnings and cyclical but steady employment. Were city conditions to deteriorate, however, the positive industry-mix effect could be insufficient to offset large, negative local-factor effects in employment and even earnings per job. The role of the terrorist attack in this scenario is discussed later.

Chart 6

Projected U.S. Job Growth in 2000-2010 versus New York City Concentration



Sources: New York State Department of Labor; U.S. Department of Labor, Bureau of Labor Statistics; Federal Reserve Bank of New York calculations.

Notes: A location quotient is a measure of local industry concentration. It is defined here as an industry's share of employment in New York City divided by its national share. For example, the securities industry accounts for 4.9 percent of jobs in New York City, but only 0.57 percent of jobs nationwide, so its location quotient is 4.90/0.57=8.6. Thus, this industry's share of New York City employment is 8.6 times the U.S. average. The dark circles represent the chief two-digit Standard Industrial Classification industries in New York City.

Average Annual Growth in Earnings and Employment
Percent

Category	Total Earnings, 1995-2000	Employment, 1995-2000	Earnings per Worker, 1995-2000	Employment, 2000-2010
National growth	7.1	2.5	4.5	1.7
Local-factor effect	-0.6	-0.5	-0.1	—
Industry-mix effect	+2.3	+0.5	+1.7	+0.3
Local growth (New York City)	8.8	2.5	6.1	—

Source: U.S. Department of Labor, Bureau of Labor Statistics.

The Financial Sector

New York City has a favorable and diverse industry mix—at least within the service sectors—yet financial services stands out as a key sector driving the local economy. The securities industry alone accounted for 5 percent of city employment in 2000 and an estimated 20 percent of city earnings. These shares were higher than they were at any point in history and roughly eight times the respective U.S. figures. Not surprisingly, then, trends in the financial sector dominate the local economic landscape. In fact, a recent study (Bram and Orr 1999) finds that cycles in the financial sector—particularly the securities industry—tend to presage cycles in the broader local economy.

Accordingly, the steep contraction in Wall Street employment over the past year—driven by the national recession, particular weakness in the financial markets, and the September 11 attack—is a matter of serious concern and appears to pose the greatest threat to the city's medium-term growth prospects. The fact that securities employment has also fallen, albeit somewhat less sharply, in the rest of the country suggests that these job losses reflect a combination of industry- and city-specific factors. Financial services will be an important sector to monitor when gauging the overall health not only of Lower Manhattan, but also of the broader regional economy. Over the longer term, the Internet and general advances in information and communications technology are likely to continue to have important effects on the broad structure of the financial services industry (see Orr and Rosen [2000] for a more complete discussion).

Financial services will be an important sector to monitor when gauging the overall health not only of Lower Manhattan, but also of the broader regional economy.

Labor Force Composition

New York City's population growth accelerated over the past decade. According to the decennial U.S. census, the city's population grew 9.4 percent during the 1990s—the strongest growth since the 1920s. Moreover, adjacent counties across the metropolitan area also saw increasingly brisk gains. Because New York City's economy tends to be a major hub of information-based industries, the metropolitan area's labor force, not surprisingly, has a much higher proportion of college graduates than the nation overall. Much of the area's highly educated population resides in the suburbs, but the city itself has attracted a very well-educated workforce. In the city proper, an estimated 30.5 percent of adults are college graduates—a share well above the national average of 25.2 percent. More important, though, is the *change* in educational attainment: in 1991, just 22.1 percent of New York City adults were college graduates, only slightly more than the national figure of 21.4 percent.[7] This indicates that New York is growing in attractiveness as a residential location for well-educated Americans. Because high concentrations of well-educated residents are beneficial to urban growth (Rauch 1993), this component of the city's population base is a positive factor for its future prosperity. In addition, the city is home to many immigrants, presumably because of its historical role as an immigration gateway. Although the effect of large stocks of immigrants on local labor markets is unclear, recent research suggests that these individ-uals tend to support a vigorous housing market (Saiz 2002). Thus, a reduction in immigration to the United States would disproportionately affect New York, if only in terms of population growth.

In sum, our empirical examination of New York City's changing industry and population mix shows the city to be attractive to those industries that benefit from its productive amenities and to a relatively highly skilled and highly educated labor force. These results are consistent with the model of a mature urban economy as well as with the rising land prices and real wages observed in the city over the past two decades.

Although the city is currently well positioned for growth over the medium term, the aftermath of September 11 could disturb this equilibrium. For example, a change in the medium-term industrial outlook is possible, but it does not appear likely at this time. A more likely factor that could affect this equilibrium is the city's policy response to the economic pressures arising from the attack. The task now faced by New York City is to restore and maintain the productive amenities and attractive environment for both firms and workers.

Policy Challenges after September 11

Although the city's industrial and population structures on September 10 were favorable for its continued prosperity, the loss of life and property on September 11 represents a journey into uncharted waters. Accordingly, we now explore certain

effects of the attack in more detail and examine whether the specific types of damage to the city's economy might jeopardize its future. In the language of the model described earlier, the attack most likely induced shifts in the household and firm curves. But how large were these shifts, and what role can policy play in restoring the city to its pre-attack equilibrium? To answer these questions, we consider the future policy implications and challenges for the city in terms of its finances, its infrastructure, and the subsequent federal aid package.

Closing the City's Budget Deficit

The September 11 attack caused a significant short-run fiscal drain on the city. Although it is difficult to isolate the attack's effect on New York City's tax revenue, city agencies estimate that the attack will have reduced revenue by 3 percent to 7 percent in fiscal year 2001-02 and by 2 percent to 6 percent in fiscal year 2002-03.[8] Although the losses represent a sizable shock to the city's revenue stream, they are considerably less severe than the losses faced by the city during its 1975 fiscal crisis. In 1975, the city had to refinance a net cash deficit of more than 25 percent of total tax revenue. The city revenue costs of the September 11 attack are much more in line with the reversal in the city and national economies in fiscal year 1990. That year, revenues from the city's four major taxes (personal income, sales, property, and general corporation) fell approximately 0.5 percent after having grown by more than 3.7 percent the previous year—a swing of 4.2 percent in revenue growth rates (Haughwout 1997). Thus, the tax revenue shortfall related to September 11 is of a magnitude similar to the effects of the city's 1990 recession.

The first policy challenge for the city . . . is to close its current budget gaps at the least possible cost to its long-run vitality.

Revenue shortfalls of this magnitude require concerted action to avoid ending the fiscal year with a deficit. In New York, deficit spending led to the severe fiscal crisis of 1975; since then, the city has been under a strict balanced budget regime. In 1990, faced with a shortfall of similar magnitude, the city raised personal income tax rates sharply to bring the budget into balance in fiscal year 1991. Today, the issue facing city officials is how to react to the current shortfall, which has been compounded by the effects of the recession and declines in equity markets. The historical pattern in New York has been to raise taxes to close budget gaps, and this is indeed the quickest way to generate cash. However, tax increases—particularly at this point in time—may be costly in the long run.

Recent research into the relationship between tax rates, bases, and revenues in four U.S. cities, including New York, indicates that these cities are very near the top of their local "revenue hills" (Haughwout et al. 2000). This means that increases in tax rates can reduce the base over time in such a way as to generate very little additional revenue in the long run. Instead, they may distort the local economies, leading to various forms of tax avoidance, especially the loss of economic activity to other jurisdictions.

Therefore, although tax increases could help close the city's current budget gap in the short run, they might well come at a substantial long-run cost, as higher taxes would make the city a less attractive location for mobile firms and households. Haughwout et al. assert that the deleterious effects of city tax increases can extend beyond the tax base by also reducing the size of the real economy and the job base. Such a reduction in the job base, unlike the temporary displacement of jobs from the September 11 attack itself, could last for the duration of the tax increases. Of course, temporary tax increases might have smaller and less durable effects.

The first policy challenge for the city, then, is to close its current budget gaps at the least possible cost to its long-term vitality. The alternative to tax increases, of course, is expenditure reductions, themselves costly. In addition, it may be viable for New York City to obtain federal and state aid to help it through this difficult period.

Rebuilding Damaged and Destroyed Infrastructure

Compounding the fiscal problems facing the city and its residents is the fact that a significant part of the damage to Lower Manhattan was sustained by public facilities, notably the Port Authority Trans-Hudson (PATH) train station in the World Trade Center. In addition, some streets and city subway lines were heavily damaged. These facilities allow Lower Manhattan to function as an employment center for the city and the region. Without ready access to Lower Manhattan, one of the city's growth engines—the financial services and related industries concentrated around Wall Street—is cut off from the rest of the region.

More generally, evidence from the academic literature indicates that the productivity of city public capital (that is, its

value to firms) is modest, but positive. Eberts (1990) and Haughwout (2002), for example, estimate output elasticities of around 0.04, meaning that a loss of 1 percent of a city's public capital stock is associated with a 0.04 percent decline in city productivity. Interestingly, household costs are estimated to be somewhat higher, yielding a total land price elasticity of about 0.12. These figures suggest that infrastructure stocks exert a significant influence on a city's attractiveness.

Initial reports indicate that New York's public facilities experienced approximately $1.4 billion in damage from the September 11 attack (Bram, Orr, and Rapaport 2002). This amount is approximately 1.5 percent of the city's preexisting infrastructure stock (Haughwout and Inman 1996), implying a 0.18 percent decline in city land values as a result of lost infrastructure, assuming that the losses are permanent.[9] The losses, however, are unlikely to be permanent.

Compounding the fiscal problems facing the city and its residents is the fact that a significant part of the damage to Lower Manhattan was sustained by public facilities, notably the PATH train station in the World Trade Center.

The importance of these facilities to New York City can be illustrated by way of a rough calculation of the value of the World Trade Center's PATH station. With the destruction of that facility on September 11, some 20,000 New Jersey–Lower Manhattan commuters were forced to find alternate means of transportation to work. For some, this required a shift to more expensive ferries crossing the Hudson River from Hoboken, New Jersey. But for the substantial number who had commuted from the PATH station in Newark, New Jersey, the quickest option was to take a commuter rail to Pennsylvania Station in New York City and then a city subway to Lower Manhattan. New Jersey Transit, the state's commuter line, estimates that daily ridership to Pennsylvania Station increased by 15,000 passengers following September 11, despite the immediate loss of jobs experienced (New Jersey Transit 2001).

For these riders, a twenty-two-minute ride from Newark to Lower Manhattan was replaced by a twenty-minute train ride to midtown plus a twenty-to-thirty-minute subway ride downtown, with a similar lengthening of the homeward commute. The additional time cost is thus approximately one hour per commuter per day. If we assume that 15,000 workers who value their commuting time at $25 per hour (half their hourly wage) are spending an additional hour per day commuting, we obtain a daily cost of $375,000, implying that the first year's loss of the World Trade Center PATH station cost nearly $100 million in lost time.[10] Thus, simply replacing the station would prevent this annual cost from becoming a perpetual loss, with a present value of nearly $2 billion.[11]

These calculations are intended to give an order-of-magnitude estimate of the value of Lower Manhattan's infrastructure systems to the city and the region. In the short run, New Jersey commuters and their employers are paying the costs of this lost infrastructure, as workers must either spend less time at work or less time with their families and friends. Although it is less certain exactly who will gain the most, rebuilding the PATH station will clearly benefit the metropolitan area's economy as a whole. Moreover, any additional improvements to Lower Manhattan's transportation linkages should further enhance the regional economy's potential.

As of this writing, the Port Authority of New York and New Jersey has publicly committed to rebuilding the PATH station, and the city has made substantial progress in repairing other damaged elements of the downtown infrastructure. Ensuring accessibility to established job centers in the region is thus a crucial policy challenge facing the city; doing so should reverse the initial negative effect of September 11 significantly.

Using Federal Compensation Effectively

Since the attack on the World Trade Center, much has been done to foster the recovery of New York City, and Lower Manhattan in particular. An outpouring of support nationwide has sought to provide spiritual and financial aid to the victims and to the city as a whole. Several blue-ribbon committees have been formed to assess the damage and monitor the progress made in rebuilding. To date, the federal government has committed $21.7 billion to the rebuilding effort. These funds will be allocated to various efforts to sustain and rebuild New York, ranging from subsidies to residential properties, to the accelerated depreciation of business plants and equipment, to the repair and reconstruction of local streets (New York City Independent Budget Office 2002).

It is therefore important to address the difficult question of the degree to which federal aid will help in the city's recovery. The city is awaiting a final accounting of the unreimbursed losses before it can know with certainty what proportion of the losses the federal aid will offset. Nevertheless, a recent agreement among federal, state, and city officials should greatly

increase the flexibility in the allocation of federal aid for rebuilding. Specifically, roughly $9 billion has been set aside by the Federal Emergency Management Agency (FEMA) and earmarked for specific functions related to site cleanup, including debris removal, restoration of damaged public facilities and equipment, and police and fire department overtime. According to the New York City Independent Budget Office (2002), this overall funding allocation is likely to be far greater than the unreimbursed cleanup losses, including the cost of restoring damaged facilities and equipment, which are now estimated to be more on the order of $3 billion. Under the recent agreement, part of the remaining $6 billion of FEMA assistance can now be redirected to help rebuild and improve the entire transportation infrastructure in Lower Manhattan, including such projects as a new PATH terminal and a new transit center that would connect to the terminal and link the downtown subway lines. The use of FEMA funds for transportation projects that go beyond the repair and restoration of damaged facilities is unprecedented.

These issues point to the third challenge facing city policymakers: to apply federal funds in the most efficient way possible to maximize their return. Overall, the federal aid package seems likely to cover the rebuilding costs of the physical facilities lost on September 11. There may even be sufficient funds, if applied judiciously, for Lower Manhattan to transform itself into an even more appealing location than it was on September 10. The proposed transportation improvements, for example, could make Lower Manhattan a more desirable place for businesses, while subsidies to residential development could enhance the neighborhood's various amenities.

Land Prices Suggest a Positive Economic Outlook

In conjunction with the model, we can gain insight into the city's longer term economic outlook by examining land prices.[12] Although these prices cannot be observed directly, the selling prices of houses, apartments, and commercial properties can serve as a rough proxy.

In the days and weeks immediately following the terrorist attack, there was widespread concern that large numbers of people would no longer want to live in New York City. As it turns out, following a brief but sharp dip in September and October, housing markets in New York City's nearby suburbs, the outer boroughs, and even Manhattan are reported to have rebounded strongly (Bram, Orr, and Rapaport 2002). As shown in Chart 3, the selling prices of existing homes throughout the metropolitan area have continued to rise, not only in absolute terms, but also relative to the national average. These results suggest that the terrorist attack thus far has had little if any negative effect on land prices, even in areas close to the World Trade Center.

Manhattan's office market, however, has been more mixed. The rental market has clearly been weak: at midyear 2002, office vacancy rates were reported to be sharply higher than they were a year earlier and rents were reported to be sharply lower. However, the buyer's market for office property continued to be characterized as strong, again suggesting underlying strength in land prices.

Going forward, the strength of the city's property markets will be an important barometer of its economy.

Thus, conditions in New York's important property markets currently appear to reflect continued strong demand. To the extent that business and household evaluations of a location's prospects are reflected in bids, New York City's land markets suggest a very positive indicator for the city. Naturally, market participants are operating with very limited information about the future. Yet it appears that, from what they know, firms and households still view New York as an attractive location. Going forward, the strength of the city's property markets will be an important barometer of its economy.

Conclusion

Economic growth in New York City for most of the past three decades has been characterized by a transformation of employment from relatively low-paying jobs to increasingly higher paying ones—that is, overall employment levels have remained roughly constant but real earnings levels have risen both in absolute terms and relative to the nation.

In light of the destruction that occurred on September 11, several downside risks to the city's economic growth prospects have arisen. The worst-case scenario includes the possibility that financial firms located in Lower Manhattan will leave the city, which could generate a cumulative downward spiral of job and income growth.

In assessing this and other potential risks to New York, we examined the city's economic position with respect to several key growth fundamentals prior to September 11. We found that New York's industrial mix, which is weighted toward high-growth sectors, along with the city's attractiveness to workers and firms, point to favorable prospects for growth. However, whether these conditions translate into actual growth going forward depends to a large degree on the city's policy response to the economic pressures arising from the terrorist attack. The main elements of this response will be the ability to manage relatively large budget deficits and rebuild New York City's damaged and destroyed infrastructure—while maintaining the productivity of the capital stock. Addressing the latter issue will require a vision of Lower Manhattan's role in the overall city economy as well as an efficient allocation of the anticipated federal aid.

Endnotes

1. By "relative," we mean in comparison to other locations within the broad capital and labor market. In practice, the nation is a sensible benchmark, as national labor supply is relatively inelastic. Capital may be more elastically supplied at the national level, but in this context we take the national supply of capital as given.

2. This conclusion is also based on the assumption that higher prices of constant-quality housing reflect growth in demand for the land on which a house sits or for durable capital in place there, and that firms and households are competitors in the land market.

3. See Harrigan and Martin (2002) and Quigley (1998) for discussions of the local benefits of industry concentrations.

4. For example, for evidence on city schools' math test scores, see New York City Department of Education (2002).

5. Because of improvements in the relative skill level of New York City employers and residents over time, we must be cautious in interpreting any income data. If the local skill-adjusted wage is higher than it is elsewhere, it can be interpreted as evidence that firms are willing to pay relatively more to workers here, implying that these workers are more productive when they work in New York.

6. See Harrigan and Martin (2002).

7. Data refer to all persons twenty-five years of age and older. City data are from the U.S. Census Bureau (1991, 1999); national data are from the U.S. Census Bureau (2002). In 1991, the city's college-educated share of the population was 22.1 percent, while the national share was 21.3 percent in 1990 (U.S. Census Bureau 1991, 2002).

8. The lower figures in the range (for both years) are from the New York City Office of the Comptroller (2001); the higher figures are May 2002 estimates from the New York City Office of Management and Budget, as reported by the U.S. General Accounting Office (2002).

9. Haughwout and Inman's data are for 1992 and should thus be viewed as illustrative of the order of magnitude of the September 11 attack.

10. The value-of-time estimate as half the gross hourly wage is standard in the urban transportation literature; see Small (1992).

11. These calculations exclude the additional out-of-pocket cost to New Jersey commuters, as these amounts are transfers to regional transportation authorities. Lost time, however, is gained by no one. Present-value calculations assume 3 percent real discount and 2 percent depreciation rates.

12. Bram, Orr, and Rapaport (2002) discuss current economic indicators since the attack in order to assess short-term prospects.

References

Blomquist, Glenn, Mark Berger, and John Hoehn. 1988. "New Estimates of Quality of Life in Urban Areas." American Economic Review 78, no. 1: 89-107.

Bram, Jason, and James Orr. 1999. "Can New York City Bank on Wall Street?" Federal Reserve Bank of New York Current Issues in Economics and Finance 5, no. 11 (July).

Bram, Jason, James Orr, and Carol Rapaport. 2002. "Measuring the Effects of the September 11 Attack on New York City." Federal Reserve Bank of New York Economic Policy Review 8, no. 2 (November): 5-20.

Eberts, Randall. 1990. "Cross-Sectional Analysis of Public Infrastructure and Regional Productivity Growth." Federal Reserve Bank of Cleveland Working Paper no. 9004.

Gyourko, Joseph, and Joseph Tracy. 1991. "The Structure of Local Public Finance and the Quality of Life." Journal of Political Economy 99, no. 4: 774-806.

Harrigan, James, and Philippe Martin. 2002. "Terrorism and the Resilience of Cities." Federal Reserve Bank of New York Economic Policy Review 8, no. 2 (November): 97-116.

Haughwout, Andrew. 1997. "After the Fall: An Examination of Fiscal Policy in New York City since 1974." In Proceedings of the Eighty-Ninth Annual Conference on Taxation, 96-105. Washington, D.C.: National Tax Association.

———. 2002. "Public Infrastructure Investments, Productivity, and Welfare in Fixed Geographic Areas." Journal of Public Economics 83, no. 3: 405-28.

Haughwout, Andrew, and Robert Inman. 1996. "State and Local Assets and Liabilities, 1972-1992." Unpublished paper, Federal Reserve Bank of New York.

Haughwout, Andrew, Robert Inman, Steven Craig, and Thomas Luce. 2000. "Local Revenue Hills: A General Equilibrium Specification with Evidence from Four U.S. Cities." NBER Working Paper no. 7603, March.

McCarthy, Jonathan, and Charles Steindel. 1997. "National and Regional Factors in the Metropolitan Economy." Federal Reserve Bank of New York Economic Policy Review 3, no. 1 (February): 5-19.

New Jersey Transit. 2001. "Call to Action: Addendum." December. Addendum to "NJ Transit's Call to Action: An Investment for the Future." September 10.

New York City Department of Education. 2002. "1999-2002 State and City Mathematics Test Results, Grades 3, 4, 5, 6, 7, and 8." Available at <http://www.nycenet.edu/daa/2002math/default.asp>.

New York City Independent Budget Office. 2002. "World Trade Center Aid: Too Much for Some Needs, Not Enough for Others?" Newsfax no. 100, May 28.

New York City Office of the Comptroller. 2001. "The Impact of the September 11 Attack on New York City's Economy and Revenues." October 4.

New York City Partnership and Chamber of Commerce. 2001. "Economic Impact Analysis of the September 11th Attack on New York City." November.

Orr, James, and Rae Rosen. 2000. "The Financial Services Sector in London and New York." In The London-New York Study, 11-24. London: The Corporation of London.

Quigley, John. 1998. "Urban Diversity and Economic Growth." Journal of Economic Perspectives 12, no. 2: 127-38.

Rauch, James. 1993. "Productivity Gains from Geographic Concentration of Human Capital: Evidence from Cities." Journal of Urban Economics 34, November: 380-400.

Roback, Jennifer. 1982. "Wages, Rents, and the Quality of Life." Journal of Political Economy 90, no. 6: 1257-78.

Saiz, Albert. 2002. "Immigration and Housing Rents in American Cities." Unpublished paper, Harvard University.

References (Continued)

Small, Kenneth. 1992. Urban Transportation Economics. New York: Harwood Academic.

U.S. Census Bureau. 1991, 1999. "New York City Housing and Vacancy Survey." Washington, D.C. Available at <http://www.census.gov/hhes/www/nychvs.html>.

———. 2002. "2001 Statistical Abstract, Table 215." Washington, D.C.

U.S. General Accounting Office. 2002. "Review of the Estimates for the Impact of the September 11, 2001, Terrorist Attacks on New York Tax Revenues." GAO-02-882R. July 26. Available at <http://www.gao.gov>.

James Harrigan and Philippe Martin

Terrorism and the Resilience of Cities

- The September 11 attacks highlighted the vulnerability of major U.S. cities to terrorism, leading some observers to question whether cities will continue to thrive despite the risk of further assaults.
- Current economic theory provides a strong basis for arguing that cities can withstand terrorist attacks and other catastrophes. The same forces that are thought to lead to the formation of cities—namely, the gains derived from the proximity of firms to markets, suppliers, and a large labor pool—will help to preserve cities in the face of an attack.
- To be sure, the ongoing threat of terrorism—a kind of "tax" on a city's firms or residents that reflects the costs of fear, higher insurance premiums, and increased security spending—may temper the economic forces that sustain cities. Nevertheless, the available evidence suggests that the size of this tax is too small to undermine the viability of New York and other large cities.

Large American cities have thrived in recent decades despite the challenges of crime, congestion, and poverty. Can terrorism jeopardize this vitality? The September 11 attacks were targeted at two major U.S. cities. Although the loss of life and the destruction were unprecedented, in one key respect the attacks were typical: terrorism in the developed world has usually been concentrated against the financial and political centers of power. The attacks on the World Trade Center and the Pentagon have forced Americans to confront the fact that to live or work in a major city is to be at greater risk of large-scale terrorism.

In this article, we consider what these risks, and the public perception of them, mean for cities in general and for the future of New York City in particular. We begin by examining the question of why cities exist at all. Only by answering this important question can we think more clearly about the long-run effects of terrorism on cities.

The economic analysis of cities is an active area of research, and we draw on that work for our study. The consensus among economists is that there are a number of plausible explanations, or models, for the existence of cities. Accordingly, we apply two economic geography models—the "labor pooling" and "core periphery" models—to our analysis of the existence of cities and terrorism's effect on them. Our use of these two different theoretical models gives a broad perspective to our study.

Although the models differ, the conclusions we draw from them are similar. Namely, the very forces that lead to city

James Harrigan is a senior economist at the Federal Reserve Bank of New York; Philippe Martin, formerly an economist at the Federal Reserve Bank of New York, is a professor of economics at the University of Paris-1.

The authors thank two anonymous referees, as well as Rohit Vanjani for excellent assistance. The views expressed are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System.

formation also enable cities to be highly resilient in the face of catastrophes such as terrorist attacks, because they create an impetus for the concentration of economic activity that is very difficult to overcome. In addition, although a continued threat of terrorism in large cities may make them a somewhat less attractive place to live and conduct business, terrorism is quite unlikely to cause massive changes in the structure of urban life. New York City in particular is likely to continue to thrive despite any ongoing threat of terrorist activity.

Why Do Cities Exist in the First Place?

A city is a dense concentration of people and economic activity. In an open society, where people and businesses are free to locate wherever they like, such concentrations can persist only if they are beneficial to the people and firms that locate there. In the language of economists, the existence of cities is part of a *market equilibrium*, where the benefits of being in a city are at least as large as the costs. Similarly, for those who do not reside in cities, the benefits of being outside a city must exceed the costs.

The study of the existence of cities in equilibrium is the province of economic geography. A general feature of economic geography models is that they include strong forces of circular causality: since there are benefits to being in cities, workers and firms move there, which increases the benefits of being in the city, causing more workers and firms to move there, and so on. An implication of circular causality is that, once formed, cities are highly robust. The large gains from being in a city mean that firms and workers are reluctant to forgo those gains even if the costs of being there rise.

What are the economic benefits of being in a city? For one, moving goods and people across space is costly in terms of time and money, and being in a city economizes on these transport costs. Transport cost benefits can be broken down further according to three forces:

- Market access—firms benefit from being close to large concentrations of consumers. For goods producers, this proximity is often a matter of economizing on shipping costs, while for many service firms (hotels, restaurants, most retailers), it is simply impossible to sell to distant consumers.
- Supplier access—producers benefit from quick access to a wide range of inputs. This access gives firms an incentive to locate near their suppliers and, since nearby suppliers are favored by producers, suppliers have an incentive to locate near their customers.
- Cost of living—consumers benefit from being near producers, since they will have less expensive access to a wide range of goods and services.

These three forces work together to create an incentive for agglomeration. A final force is the fact that workers have to live near their jobs. Because workers are also consumers, by coming together in a single location, firms can sell to each other as suppliers and to each other's employees as consumers. Economists have worked hard to tidy up the loose ends of this argument, but the intuition is clear enough: when there are transportation costs, agglomeration is an equilibrium.

An alternative, but complementary, reason for cities is that their existence makes it easier for workers and firms to find each other. With a large pool of workers to choose from, employers can quickly hire the people they need. Similarly, with many potential employers, workers can find a job quickly.

> *In the language of economists, the existence of cities is part of a* market equilibrium*, where the benefits of being in a city are at least as large as the costs.*

If firms' demand for labor fluctuates over time, this labor market pooling effect can be particularly valuable for both firms and workers: when a firm lays off workers, those workers can find new jobs more easily if they live in a big city than if they live in an isolated company town; conversely, when a firm is booming, it can more easily lure new workers if a large pool of job seekers exists.

An appealing aspect of the labor market pooling motivation for cities is that it does not rely on transport costs, except to the extent that workers need to live near their jobs. Even as economic activity becomes more "weightless" and easy to transmit over space at low cost, the labor market pooling motivation remains undimmed.

To see the possible relevance of labor market pooling to New York City, consider the financial services sector—the backbone of the city's economy and tax base. Financial services have evolved to the point where face-to-face interaction between customers and producers is almost completely unnecessary for many transactions, and access to a wide range of produced inputs is equally irrelevant.[1] Yet financial services firms remain in New York because that is where their workers

are, and their workers are there because the firms are there. As long as some work needs to be done by groups of workers located in the same place, economic activity cannot be completely dispersed, and agglomeration will persist even if final sales can be transported at zero cost.[2]

However, along with the benefits of being in a city come costs. The most obvious is congestion: cities are crowded and urban land prices are high. Another drawback for firms locating in a city is the heightened competition that arises from locating in an area where there are already many firms selling similar products. In equilibrium, firms and workers balance the benefits and costs of locating in a city, and the size and durability of cities depends on this trade-off.

Once a city is formed, the cost-benefit calculations that led to its creation may no longer be relevant. This is because the existence of the city is in itself a reason for its persistence. In particular, the level of costs that could lead a city to erode is generally far higher than the costs sufficient to discourage city formation in the first place. In short, there is an *agglomeration rent*: once workers and firms are in a city, costs and conditions can deteriorate substantially without tempting them to leave.[3]

How Does Terrorism Affect a City?

The impact of September 11 on New York City was not limited to the death and destruction in Lower Manhattan. The cost of doing business in New York has risen, even for workers and firms quite distant from the disaster site.

Some of these costs include:

- slower and less efficient transportation, because of the physical disruption of transport links and the delays caused by increased security,
- greater security spending,
- higher insurance premiums,
- the emotional toll on workers who fear future attacks.

In brief, the ongoing costs imposed by September 11 are significant. Indeed, some of these costs are now being borne in many large U.S. cities, as citizens have seen the damage inflicted on New York and are taking cautionary steps in response. The effect of these increased costs is to make cities less attractive—an effect that will last as long as the threat of terrorism persists.

Having established that cities exist because of the decisions of millions of independent workers and businesses, we return to our central question: what are the implications of terrorism for the future of cities? To answer this question, we turn to two specific models of the existence of cities, and examine the effects of terrorism in these models. The models illustrate the robustness of cities: they suggest that even in the face of the large and continuing costs associated with terrorism, cities are likely to continue to thrive.

The Labor Pooling Model

For our first model, we use a modified version of a simple model of labor market pooling introduced in Krugman (1991). The details of the model are given in the appendix, but the economic intuition is explained here.

The ingredients of the model are firms, workers, and locations. Firms choose their location to maximize profits,

Firms want to locate where other firms are to reduce fluctuations in wages.

while workers move to where wages are highest. In addition to labor, firms use land to produce their output, and land is in fixed and equal supply in each of two locations, A and B.

Firms face uncertainty, and dealing with this uncertainty is an important part of their business plans. The uncertainty can reflect fluctuations in demand or in productivity, and it is important to theorize that the uncertainty is imperfectly correlated across firms: when some firms face high demand and/or productivity, other firms face low demand and/or productivity.

Because of the fluctuations facing firms, wages will fluctuate in each location: when many firms face high demand or have high productivity, wages will be bid up, and wages will fall when labor demand is low. But the variability in wages in a given location will be lower the more firms are in that location. This relationship follows from the law of averages: firm-specific positive and negative fluctuations are more likely to cancel each other out if there are many firms. In this case, firms that benefit from high demand or high productivity will be able to hire new workers without bidding up wages. This would not be true if there are few firms in the same location, or if firms' shocks are correlated. Hence, a larger number of firms and more diversity in a location will raise average profits in that location. In brief, firms want to locate where other firms are to reduce fluctuations in wages.

Land is in fixed supply in each location, and since each firm needs land as well as labor, the price of land is higher where there are more firms. This introduces a trade-off for firms: by locating where other firms are concentrated, firms benefit from a larger labor supply, but they must pay higher rents.

Turning to the decisions of workers, we assume that they care only about average wages.[4] As a result, if average wages are the same in two locations, workers will be indifferent about where they live. Since wages are determined by supply and demand, wages will be equal in each location if the share of workers equals the share of firms in each location. We illustrate this relationship in Exhibit 1: let λ_F be the share of firms in location A, and λ_L be the share of workers in location A. The 45 degree line is where $\lambda_F = \lambda_L$ so that wages are equalized, and we label the 45 degree line *WW* to indicate this equality. Above *WW*, $\lambda_F > \lambda_L$, so the share of firms in A is higher than the share of workers; this means that wages are higher in A than in B. This would lead workers to migrate from B to A, and the opposite holds for points below *WW*. These potential migrations are indicated by the arrows in the exhibit. Thus, with free migration, *WW* is the labor market equilibrium condition.

The location choice for firms is a bit more complicated, since it is based on three considerations: the average level of wages, the variability of wages, and the cost of land. In equilibrium, it must be the case that average profits are the same in each location, otherwise firms would move to the more profitable location. This trade-off is illustrated in Exhibit 2, which, like Exhibit 1, has location A's share of firms, λ_F, and share of workers, λ_L, on the two axes. As we detail in the appendix, the $\pi\pi$ curve shows the combinations of λ_F and λ_L that are consistent with equal average profits in each location. Starting at the symmetric point where there are equal numbers of workers in each location ($\lambda_F = \lambda_L = 1/2$), we note that there are two possibilities:

- If the labor pooling motive is weak and/or if the price of land responds strongly to the number of firms, the $\pi\pi$ curve has a slope of less than 1, illustrated as $\pi\pi_1$.
- If the labor pooling motive is strong and/or if the price of land does not respond much to the number of firms, the $\pi\pi$ curve has a slope that is initially greater than 1, but then flattens out, as illustrated by $\pi\pi_2$.

In either case, points below $\pi\pi$ are combinations of λ_F and λ_L such that profits are higher in A than in B, so that firms will move from B to A, with the opposite holding for points above *WW*. These potential movements of firms are indicated by the arrows in Exhibit 2.

With the labor market and firm location equilibrium conditions established, we can now determine the overall equilibrium location of workers and firms. This is where the *WW* and $\pi\pi$ curves intersect. The case where $\pi\pi$ is fairly flat at $\lambda_F = \lambda_L = 1/2$ is not very relevant, since it implies that the two locations are symmetric and no city emerges, so instead we focus on the case where $\pi\pi$ is steep at $\lambda_F = \lambda_L = 1/2$ (Exhibit 3). There are two equilibria in this case, but the symmetric one at $\lambda_F = \lambda_L = 1/2$ is unstable: any slight increase in λ_F or λ_L away from this point will be self-reinforcing, as indicated by the direction of the arrows, and the economy will converge to the asymmetric equilibrium where

Exhibit 1
The Location of Workers

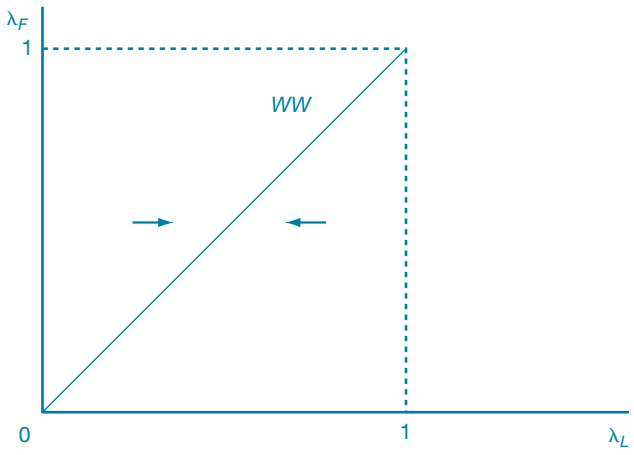


Exhibit 2
The No Agglomeration Case

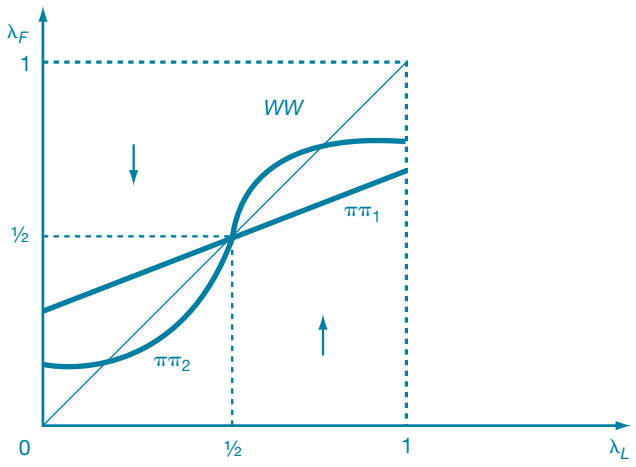
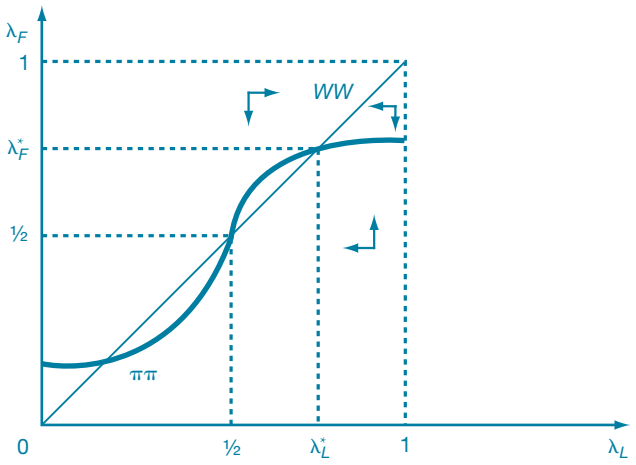


Exhibit 3
The Agglomeration Case



Note: We denote the equilibrium location of firms and workers as λ_F^* and λ_L^*, respectively.

one location is larger than the other ($\lambda_F^* = \lambda_L^* = 1/2$). At this equilibrium, location A (the city) has high land prices and firms benefit from labor pooling, while firms in location B (the nonurban location) forgo the benefits of labor pooling in exchange for lower land costs.[5]

Thus, in the labor pooling model, cities arise because the benefits to firms of a large and stable labor supply balance the high rents that must be paid in crowded locations.

The Effects of Terrorism in the Labor Pooling Model

As we have observed, terrorism targeted against large cities has two effects: the loss of life and the destruction associated with the attacks, and the costs of coping with any ongoing terrorist activity.

Perhaps surprisingly, a onetime incident in itself has no long-run effect on the size of cities in the labor pooling model. This is because even massive physical damage has no permanent effect on the underlying forces leading to the concentration of economic activity in a single location. Evidence to support our view is offered by Davis and Weinstein (forthcoming), who study the impact of the U.S. bombing of Japanese cities in 1944-45. They show that the damage to cities was both huge and very uneven, with some cities (such as Tokyo and Hiroshima) deeply affected and others relatively unaffected. However, fifty years later, the relative size of cities in Japan was the same as it was before the war; accordingly, the forces leading to city formation were left intact after the bombing stopped.

Nevertheless, *ongoing* costs can change the trade-off between concentration and dispersion, potentially leading to changes in the equilibrium amount of concentration. In the labor pooling model, terrorism in cities can directly affect firms through an increase in the cost of doing business, and affect workers through an increase in the perceived risk of living and working there. Because the decisions of firms determine local labor demand, any direct effect of terrorism on firms has an indirect effect on workers. Similarly, the direct effect of terrorism on workers' location decisions has an effect on firms because it influences local labor supply.

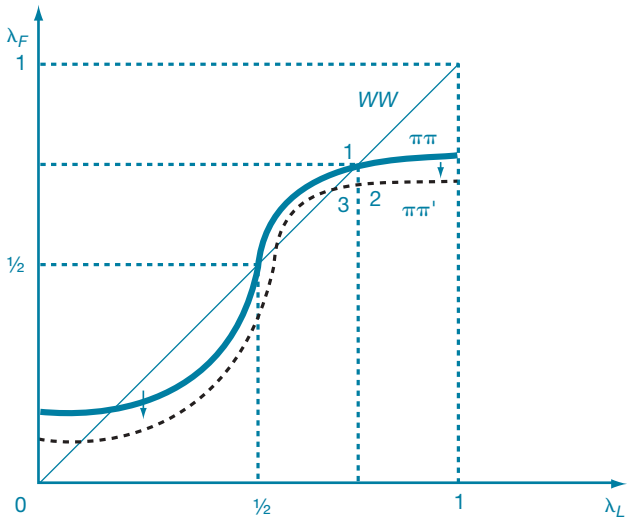
We first interpret the ongoing costs of terrorism as a cost that must be borne by firms that locate in the city. These costs include higher insurance premiums, direct spending on increased security, and reduced productivity associated with security-induced delays. We refer to these costs as a *terror tax*: the ongoing cost of doing business in a city threatened by terrorism.[6] Unlike a typical tax, this tax detracts from firm profits and worker income without funding improvements in infrastructure or services. Firms can avoid the terror tax by relocating to a smaller city, and it is this trade-off between security and agglomeration benefits that we analyze.

Terrorism in cities can directly affect firms through an increase in the cost of doing business, and affect workers through an increase in the perceived risk of living and working there.

As Exhibit 4 shows, the initial equilibrium is at point 1, and terrorism shifts down the $\pi\pi$ curve to $\pi\pi'$. Firms leave the city to avoid the costs related to terrorism, wages in the city fall in response to the fall in labor demand, and the equilibrium moves to point 2. At that point, however, wages are lower in the city so workers leave as well. The new equilibrium is at point 3, with wages and profits again equalized between locations, but with the city somewhat smaller. Profits decrease in both locations: in the city because of lower agglomeration benefits and the direct costs of terrorism, in the nonurban area because land rents are bid up. Hence, even though the direct effect of

Exhibit 4

The Effects of Terrorism on Firms in Large Cities



terrorism is localized, the indirect economic effect is ubiquitous because of firm and worker mobility.

However, agglomeration forces are still present, and location A remains big despite the terror tax, which causes some firms to leave location A for location B. The fact that gains to agglomeration still exist implies that one location is bound to have more firms and workers. Another way to say this is that an agglomeration rent exists in the city: even if the profits of firms in the large city are hit by a permanent increase in costs, most firms will want to stay to benefit from labor pooling. However, if the increase in costs is large enough, and the consequent shift of the $\pi\pi$ curve is large enough, then the terror tax will cause so many firms and workers to leave the city in location A that the benefits of labor pooling will dry up, and the city will no longer have an economic rationale.

To illustrate this point, we construct a simple numerical example. We choose parameters such that in the initial situation, with no terrorism threat, some stylized facts in urban economics are replicated (see appendix). The labor pooling gains are large enough that the equilibrium share of firms and workers ($\lambda_F = \lambda_L$) in the city is 80.5 percent, which is roughly the urbanization rate in the United States. The share of labor income in GDP is two-thirds and the land rent in the city is ten times the land rent in the nonurban area.[7] This ratio in land prices is a measure of the agglomeration rent, or extra profit to firms from being in the city: it reveals that firms are willing to pay rents that are ten times higher in the city in order to benefit from the presence of other workers and firms.

We look at different scenarios of permanent increases in the costs of doing business in the city. An important point is that these are not the costs of fighting terrorism at the federal level, which presumably are financed by all agents in the economy, but are the added costs of doing business in the city that do not affect firms in the nonurban areas.

We find that a permanent 1 percent terror tax for all firms in the city has little effect: the share of workers and firms in the city falls from 80.5 percent to 80 percent.[8] The main adjustment comes in the form of a decrease in the ratio of land rent in the city to the nonurban area, from 10 to 7.5. This adjustment in prices is clearly a stabilizing factor for the city as it retains firms. Hence, the agglomeration rent decreases but it remains very large. A 5 percent terror tax has a more pronounced effect, as $\lambda_F = \lambda_L$ decreases to 76.3 percent and the ratio of land prices falls to 6. In the numerical example, a 7 percent terror tax causes the city to cease to exist, as firms leave the city followed by workers. At that level, the terror tax is large enough to offset the gains from labor pooling in equilibrium, and firms and workers disperse across space to avoid the tax. (Nevertheless, we argue below that such a large increase in costs that would fall entirely on the city is very unlikely.)

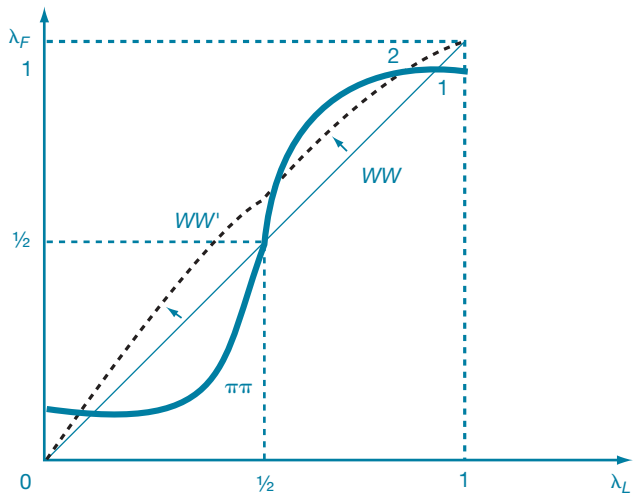
Thus far, we have assumed that the effect of a terrorist threat on cities operates through the increased costs that firms have to pay. But another effect of terrorism—and perhaps the one most desired by the perpetrators—is fear. Fear of terrorism may cause more than sleepless nights if people act on their fears by fleeing the city for locations perceived to be safer. Accordingly, it is reasonable to examine the effect that such fear has on cities in equilibrium.

Even if the profits of firms in the large city are hit by a permanent increase in costs, most firms will want to stay to benefit from labor pooling.

We model fear by supposing that workers are no longer indifferent between locations as long as wages are the same: at equal wages, they would rather be away from the city. Put differently, fearful workers demand a wage premium for working in cities because fear reduces the utility of living in the city.[9] Contrary to the case where the direct effect of terrorism is on firms, here the $\pi\pi$ curve is not affected but *WW* is, and shifts to the left (Exhibit 5). The equilibrium moves from point 1 to point 2 as both workers and firms leave the city. Profits in

Exhibit 5

The Effects of a Perceived Decrease in Safety in Large Cities



both locations decrease because of the increase in wage costs for the firms in the city and the reduction in agglomeration gains. As in the case where firms pay a terror tax, the city becomes smaller. However, if the agglomeration gains are sufficiently large, the city does not disappear, because labor pooling implies that firms are willing to pay a wage premium to retain workers in a single location, the city.

As we did with the terror tax, we construct a numerical example to illustrate the equilibrium effects of fear. Using the same parameters as before, we look first at what happens if the utility of workers living in the city is decreased by a fear factor equivalent to 1 percent of the wage.[10] The effects are found to be small: the share of workers in the city falls from 80.5 percent to 79.9 percent and the share of firms to 80 percent.[11] A 5 percent decrease in urban utility is found to have a larger effect: 76 percent of firms remain in the city and the share of workers falls to 75 percent. The city unravels abruptly when the fear factor rises to 7 percent of the wage: this exhausts the benefit to firms from labor pooling, and the economic rationale for the city is eliminated.

These numerical examples reveal a threshold effect of terrorism on the city. When the direct effect of terrorism on either the costs to firms or the peace of mind of workers is small enough, the equilibrium effect on city size is very small. But beyond a certain point (7 percent in both examples), the direct costs of terrorism have a catastrophic effect by wiping out the gains to labor pooling in the model, and the city loses its economic rationale.

Our conclusion from the labor pooling model is that the effect of terrorism on cities is discontinuous. Any ongoing costs of coping with terrorism, however, can lead to a limited reduction in the size of the city. The worst-case scenario—where the urban landscape ceases to exist—appears abruptly but only for large, permanent, increased costs of doing business in the city or for large increases in the perceived risk of living in the city.

The Core Periphery Model

The labor pooling model abstracts from transportation costs, except for the implicit assumption that workers need to live close to their jobs. But moving goods across space is costly, and transportation costs increase with distance. These facts offer an alternative explanation for the formation of cities that is independent of the labor pooling model. In the case of transport costs, businesses want to be near their customers (which may also be other businesses), and workers need to live near their jobs. Because workers are also consumers, by coming together in a single location, firms can sell to each other as suppliers and to their own and each other's employees as consumers.

To account for this interrelationship, we now apply a second model to the effects of terrorism on cities. The model we use, known as the core periphery model, is based on Fujita, Krugman, and Venables (1999). As with the labor pooling model, we begin with two *a priori* symmetric regions, and study the forces that might lead to a concentration of activity in one of the two regions (the technical details of the model can be found in the appendix).

Consider that each of two regions, A and B, is populated by an equal number of immobile workers, who we can think of as farmers.[12] There are two sectors: agriculture and manufacturing. The manufacturing sector has many firms that produce different goods and is subject to increasing returns, and manufactured goods are costly to transport. In particular, firms are assumed to incur a fixed cost for every production location, which gives them an incentive to concentrate production in a single place. Each manufacturing firm wants to sell to customers in both locations, but because of transport costs, firms will want to concentrate production in the larger market, so as to minimize their total transport cost bill. Aside from farmers, who are unable to leave their birthplace, there are manufacturing workers who are free to migrate in search of the highest real wages.

The model has two forces that encourage agglomeration:

- Market access—firms want to locate in the larger market so they can service customers without paying transport costs.
- Cost of living—if nominal wages are equal in both locations, workers would rather live where there are more firms because the price index, which includes transport costs, is lower.[13]

There is also one force opposing agglomeration:

- Competition—this force becomes more intense as more firms locate in a region, so profits become lower. Firms want to avoid competition, which gives them an incentive to locate in the region with fewer firms.

In equilibrium, the balance of these three forces depends on the detailed parameters of the model: agglomeration may be the only stable equilibrium, or it may not be an equilibrium at all. To see how the model works, it is helpful to begin by thinking about extreme cases. When it comes to very low transport costs, agglomeration is irrelevant and there is no reason for cities to exist at all, so half of all economic activity will be in each location. With very high transport costs, the benefits to firms of concentrating production in a single location are outweighed by the prohibitive costs of selling in distant markets; as a result, firms produce in both locations and there is no agglomeration.

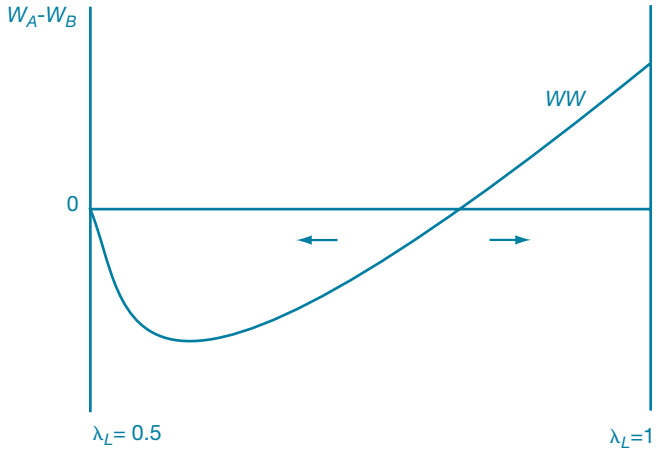
Because workers are also consumers, by coming together in a single location, firms can sell to each other as suppliers and to their own and each other's employees as consumers.

For intermediate levels of transport costs, which is the relevant real-world situation, there are two possibilities. The simplest scenario is the one where a city is inevitable: it will always pay for a worker to move to the location with more workers. Starting from symmetry, if one intrepid worker moves from location B to A, his wages will be higher, setting in motion a circular process that leads to a core periphery equilibrium, where all manufacturing is located in region A and where manufacturing wages are highest.

For our purposes, however, a second scenario is more applicable. Here, there are two locally stable equilibria, one symmetric (with half the manufacturers located in each region) and one core periphery. This scenario is illustrated in Exhibit 6, which shows λ_L as before the share of mobile workers in location A along the horizontal axis and the difference in real wages between locations on the vertical axis. The *WW* curve depicts the relationship between the real wage gap, $w_A - w_B$, and λ_L, and there are two locally stable equilibria, at $\lambda_L = 1/2$ (half of the mobile workers in each location) and at $\lambda_L = 1$ (all mobile workers in location A). To see that these equilibria are locally stable, imagine a worker in either equilibrium who decides to move to the other city: that worker will experience a drop in wages and will return home, as indicated by the direction of the arrows.[14]

The model is silent about which locally stable equilibria the economy is likely to end up in, but the core periphery equilibria at $\lambda_L = 1$ seems to be the relevant scenario for a world in which cities of unequal size coexist. Although the model is very simple, the core periphery equilibrium does have features that suggest the real world. First, there is circular causation in city formation: firms and workers want to be where other firms and workers are located. Second, wages are higher in the city, since $w_A - w_B > 0$ at $\lambda_L = 1$. Third, some economic activity (manufacturing) is concentrated in a city while other activity is spread more evenly across space.

Exhibit 6
Wages and Agglomeration



The Effects of Terrorism in the Core Periphery Model

The consequences of a onetime terrorist attack on a city in the core periphery model are similar to the effect in the labor pooling model: as long as the attack does not increase the ongoing costs of doing business in the city, there will be no long-run effect. The reason is that the underlying forces leading to city formation (a desire by firms to economize on transport costs) are unaffected by physical destruction of the city.[15]

A continuing terrorist threat that imposes ongoing costs, however, may have different consequences. As we did in the labor pooling model, we model the effect of terrorism in the city as a terror tax borne by firms. This shifts the *WW* curve down. Somewhat surprisingly, the equilibrium effect of terrorism on city size is either zero or enormous. For relatively small shifts in the *WW* curve, wages fall but remain higher in the city and there is no migration. This reaction illustrates the existence of an agglomeration rent: in the initial equilibrium, workers earn a premium for living in a city, and there is room for this premium to be eroded without causing any change in the equilibrium.

However, if the shift in the *WW* curve is large enough so that wages are lower in the city when the city exists ($w_A - w_B < 0$ at $\lambda_L = 1$), workers will start to leave the city, and the city will begin a vicious circle: as more workers and firms leave the city, the wage disadvantage persists, and the circular causation process goes into reverse (Exhibit 7). The new equilibrium is at $\lambda_L = 0.5$, where firms do not benefit from agglomeration but (by assumption) face no differential risk of terrorism because they are located in a city.

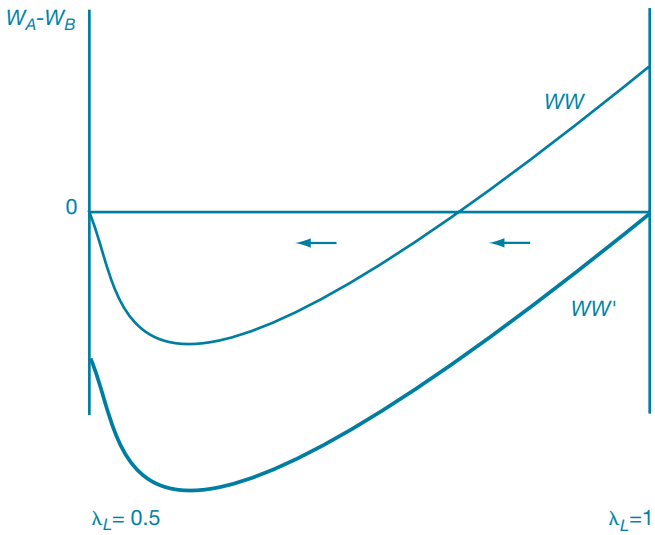
A striking aspect of this scenario is that even if the terrorist threat recedes, so that the *WW* curve returns to its original position, the dispersed equilibrium will persist. The reason is that, as noted above, there are two locally stable equilibria and the economy will tend to stay wherever it finds itself. Once the core periphery equilibrium is destroyed, it will not reappear. In such a case, the economy would be in a worse equilibrium than the one in which it began, with the efficiency gains of agglomeration permanently lost.

To offer a sense of the likelihood of the two possible scenarios, we present a numerical example. We choose parameters that imply markups of price over cost of about 25 percent, a share of goods in consumption of about 40 percent, and a level of transport costs (70 percent of sales value) that implies that there are two locally stable equilibria. This level of transport costs seems high at first glance, but it is more reasonable when nontradable goods such as most services are taken into account.

Beginning from a situation where the economy is in a core periphery equilibrium, we find that a terror tax of 1 percent or 5 percent of total costs has no effect on agglomeration: workers in the city are unhappy compared with the days before September 11, but are not unhappy enough to abandon the city. The critical value of the terror tax is 6.3 percent: if firms have to pay more than this amount to counter the effects of the ongoing threat of terrorism, they can no longer offer workers a wage premium for living in the city and workers will begin to move away, setting the vicious circle in motion. Interestingly, this critical level of the terror tax is close to the critical level of 7 percent found in the numerical examples of the labor pooling model.

In general, the vicious-circle scenario for the city is more likely to occur if the initial real wage advantage of residents is small, so that even a fairly small increase in costs in the city can tip the balance away from it as a place to work, starting the reverse cumulative process. Although this scenario seems extreme and is not likely to apply to New York City as a whole, it does make the general point that short-term disruptions to the benefits of agglomeration may have very long-term consequences.

Exhibit 7
The Effects on the City When the Terror Tax Is Large



Cities Are Unlikely to Disappear in the Face of Terrorism

We have presented two models in which an ongoing terrorist threat might jeopardize the existence of cities. The numerical examples used in the models, which we chose to reflect some of the relevant features of the U.S. economy, suggest that an ongoing terror tax (that is, higher costs paid by urban firms and residents that have no economic return) in the range of 5 percent to 10 percent might be enough to dissipate the economic rationale for the continuing existence of cities.

We should note that, despite their overall usefulness as measures of the vulnerability of New York or other cities, our numerical simulations probably understate the resilience of urban life substantially for the simple reason that each model examines just one motive for agglomeration in isolation. And although it is not possible to carefully analyze together all of the economic reasons why cities exist, it seems likely that cities are in fact much more robust than our models—which consider only one force at a time—suggest.

It is also worth noting that the level of terror tax that might cause cities to decline, as identified in our models, is far higher than anything that the war on terrorism thus far suggests is plausible. First, most of the costs discussed elsewhere in this volume (Hobijn 2002) are borne by the nation as a whole and are not disproportionately burdensome to New York or other cities. As a result, they are not an urban-specific terror tax of the kind that we analyze.

Nevertheless, one element of costs that does closely match our theoretical notion of a terror tax is insurance premiums, which can be expected to be disproportionately higher in New York as long as the threat of terrorism persists. The midyear outlook, published by Standard and Poor's in May 2002, reports that since September 11 "large commercial property insurance is commonly seeing rate increases of 50 percent-60 percent, with some policies exacting 200 percent-300 percent, along with more restrictive terms and conditions." These are large increases, and in New York some firms cannot buy insurance against the risk of terrorism at any price (Standard and Poor's Insurance 2002). However, insurance premiums do not constitute a large share of total costs. In a 1995 survey by the consulting firm Tillinghast-Towers Perrin, insurance premiums were reported to be 0.25 percent of the revenues of large firms (with more than $100 million in revenues). The amount is likely to be larger for smaller firms. For small and large firms combined, the burden of higher insurance premiums is estimated to be between 0.25 percent and 0.50 percent of firms' total costs. Yet even this small amount overestimates the urban insurance burden, because it does not measure the differential increase in costs between New York and the rest of the country.

Another element of costs specific to New York City is the ongoing disruption to transportation as a result of security delays and the destruction of the World Trade Center's Port Authority Trans-Hudson (PATH) train station. There is no way to estimate the costs of ongoing security-related delays at bridges and tunnels and the like, but anecdotal evidence suggests that the costs are not large. Bram, Haughwout, and Orr (2002) calculate that the time cost of a longer commute for former users of that PATH station is approximately $100 million per year, which is just 0.1 percent of gross city product.

Another way to think about the magnitude of the terror tax is to consider the actual dollar costs of September 11. Bram, Orr, and Rapaport (2002) estimate the value of wage losses associated with September 11 to be up to $6.4 billion and the value of physical destruction to be $21.6 billion. The size of New York's economy in 2001 is estimated to be between $381 billion (by the consulting firm DRI-WEFA) and $461 billion (by the New York City Office of the Comptroller). Assuming a 15 percent real rate of return on the destroyed capital stock (15 percent is roughly the aggregate rate of return on physical capital in the United States), the economic damage from the attack was on the order of 1.5 percent to 2.5 percent of the city's annual economic output. Although smaller than the calculations from our numerical examples, 2.5 percent is not a small number, and it might conceivably be large enough to threaten the future of New York—*if* the burden was ongoing and was not offset by insurance and federal assistance. However, as Bram, Orr, and Rapaport point out, most of the financial cost of September 11 will be covered either by private insurance or by federal assistance. Thus, the viability of urban life would be threatened only if *all* of the following conditions existed: firms were unable to obtain private insurance, the

One element of costs that does closely match our theoretical notion of a terror tax is insurance premiums, which can be expected to be disproportionately higher in New York as long as the threat of terrorism persists.

nation offered no financial assistance in the event of an attack, an attack of the destructiveness of September 11 was expected to occur every year, and the balance of forces that sustains agglomeration is even more fragile than our numerical examples suggest. It is safe to say that such a scenario is not likely.

We reiterate that our theoretical analysis of the terror tax suggests that only costs that are ongoing and specific to cities can diminish the powerful economic rationale for urban agglomeration. All available evidence suggests that the current size of the terror tax is small and is surely not large enough to threaten the viability of New York or other large cities. Moreover, the logic and the numerical simulations of the models examined suggest that in the most likely case, where agglomeration gains are sufficiently large, the effect on the size of the city will be quantitatively very small. By construction, in the core periphery model, there is no effect at all. The labor pooling model suggests that a terror tax on the order of magnitude given above would have very little negative effect on the size of the city and that the adjustment would take place through a decrease in the difference in land rents between the city and the nonurban areas. This small effect is not surprising given the circular causality mechanism at work in the model: the very fact that most workers and firms remain in the city is the source of the benefit of locating in the city. This also reflects the very nonlinear effect of terrorism on cities. It either eliminates the benefits of agglomeration or has no effect on them. Our evidence points to the latter conclusion—that terrorism will have very little effect on the size of cities.

Is There a Role for Policy?

From the standpoint of the health of cities, our analysis suggests that the best thing that government can do is to defeat terrorism, so that cities are not targets and do not have to bear its ongoing costs. That having been said, our two models suggest that, absent an expensive ongoing threat, cities will recover from terrorist attacks because physical damage does not affect the balance of forces that leads to agglomeration. If the threat of terrorism recedes, the city will return to its former size without government intervention. But in the unlikely event that terrorism does persist with very large costs, this balance of forces may be threatened, and the health of cities may be at risk. Then, and according to the logic of the models, a modest subsidy to economic activity in the city would be enough to buffer the city against the threat of firms and workers leaving. In an application of the precautionary principle, some may reasonably argue in favor of such a small subsidy as insurance against a very low-probability event.

Nevertheless, although both of our models suggest that an urban-specific subsidy may offset any negative effects of the terror tax, such a policy may not be needed. As our study indicates, it is unlikely that New York City or other cities will require subsidization to prevent them from declining in the face of potential terrorism.

Conclusion

Some observers have expressed concern that the September 11 attacks jeopardize both the present and future economic health of New York, Washington, and other major cities. In this article, we argue that the vitality of cities is unlikely to be diminished by the threat of terrorism. Rather, the forces that lead to city formation also enable cities to be highly resilient in the face of catastrophes such as terrorist attacks, because they constitute a force for agglomeration that is very difficult to overcome.

The two theoretical models that we examine suggest that the physical damage sustained from a single terrorist attack, despite its scale, is unlikely to represent a threat to the continued vitality of large cities. The effects of *ongoing* terrorist activity, however, are potentially greater, because the costs of dealing with terrorism can be viewed as a type of tax that may alter the balance of economic forces that leads to city formation. Although such a change in the attractiveness of doing business in cities may have a very modest effect, or no effect at all, there is at least the possibility that it could lead to a reversal of the circular causation that sustains the existence of cities.

Our analysis also suggests that cities in general, and New York City in particular, are highly unlikely to decline in the face of even a sustained terrorist campaign. Finally, our look at the data suggests that even a pessimistic view of the magnitude of the costs currently facing New Yorkers leads to the conclusion that the city's economic vitality is not at risk.

Appendix

In this appendix, we develop the math behind the models discussed.

The Labor Pooling Model

We use a modified version of a small model of labor market pooling that is introduced in Krugman (1991). We add to Krugman's analysis a market for land as well as shocks and uncertainty, specific to the threat of terrorism, that are concentrated in cities. In this model, the force behind agglomeration of economic activities in cities is very close to the Marshallian theory that a pooled market benefits both workers and firms. Uncertainty at the firm level plays a major role in this theory, and this type of uncertainty makes cities attractive to firms and workers.

The Basic Model

Suppose there are two locations, A and B; a total of N mobile firms; L mobile workers; and a given stock F of land in each region. n_A and n_B firms are located, respectively, in A and B. To simplify the scenario, we assume that firms, presumably because of the existence of high fixed costs, cannot have production sites in both locations. Firms are represented by a revenue function in which labor and land are the two arguments. As in Krugman (1991), we assume that the revenue function is the same in both locations, is quadratic, and that there are firm-specific shocks to the marginal product of labor:

$$(A1) \qquad R_{ij} = (\beta + \varepsilon_i)L_{ij} - \frac{\gamma}{2}L_{ij}^2 + \delta F_{ij} - \frac{\mu}{2}F_{ij}^2 - cN$$

$$i \in [1, N], \; j = A, B,$$

where $\beta, \gamma, \delta, \mu$, and c are positive parameters; L_i is the number of workers employed by the firm; and F_i is land used by firm i. ε_i is a firm-specific shock distributed normally $(0, \sigma_\varepsilon^2)$. The last term in equation A1 reflects the negative impact of competition on revenues and is not different across locations. The shocks are uncorrelated across firms. Both the labor market and the market for land are competitive and firms set the marginal products of the two factors equal to their price, respectively, w and r.

$$(A2) \qquad w_j = \beta + \varepsilon_i - \gamma L_{ij}$$

$$r_j = \delta - \mu F_{ij} \quad i \in [1, N], \; j = A, B.$$

This implies a labor demand and a demand for land from each firm:

$$(A3) \qquad L_{ij} = \frac{\beta + \varepsilon_i - w}{\gamma}$$

$$F_{ij} = \frac{\delta - r}{\mu} \quad i \in [1, N], \; j = A, B.$$

For each location, both factor markets clear:

$$(A4) \qquad \sum_{i=1}^{nj} L_{ij} = L_j \quad \sum_{i=1}^{nj} F_{ij} = F \quad i \in [1, N], \; j = A, B,$$

so that factor prices are given by:

$$(A5) \qquad w_j = \beta - \frac{\gamma L_j}{n_j} + \frac{\sum_{i=1}^{n_j} \varepsilon_i}{n_j}$$

$$r_j = \delta - \mu \frac{F}{n_j} \quad i \in [1, N], \; j = A, B.$$

Hence, the expected wage in a region depends positively on the number of firms in the region, but the variance of the wage depends negatively on that same number:

$$(A6) \qquad E(w_j) = \beta - \frac{\gamma L_j}{n_j} \quad var(w_j) = \frac{\sigma_\varepsilon^2}{n_j} \quad j = A, B.$$

The reason is that as shocks are uncorrelated across firms, the effect of a positive labor productivity shock on one firm on the equilibrium wage rate will be compensated by other firms' negative shocks, the more so that more firms are located in that location. This is in essence a specific version of the Marshallian argument of labor pooling. The covariance of the wage with firms' specific shocks also depends negatively on the number of firms: $covar(w_j) = \frac{\sigma_\varepsilon^2}{n_j} \quad j = A, B$.

The equilibrium demand for labor from a specific firm depends positively on its shock and negatively on other firms' shocks:

$$(A7) \qquad L_{ij} = \frac{1}{\gamma}\left(\varepsilon_i + \frac{\gamma L_j}{n_j} - \frac{\sum_{i=1}^{n_j} \varepsilon_i}{n_j}\right) j = A, B.$$

Using the definition of profits as revenues minus factor costs, and equations A5 and A7, we can obtain the expected profits in each location:

$$(A8) \qquad E\pi_j = \frac{\gamma}{2}E(L_{ij})^2 + \mu\frac{F^2}{2n_j^2} - tN$$
$$= \frac{\gamma L_j^2}{2n_j^2} + \frac{\mu F^2}{2n_j^2} + \frac{(n-1)}{2n\gamma}\sigma_\varepsilon^2 - tN \quad j = A, B.$$

Note that the expected profits in a location depend positively on the number of workers in that location. This is simply because the wage rate decreases in the supply of workers. The expected profit in a location has a more ambiguous relation to the number of firms in that location. On the one hand, more firms implies a higher demand for factors that bid up their price and lower profits (the first two arguments of equation A8). On the other hand, a larger number of firms makes wage rates more stable, which in turn increases expected profits, as profits are concave in the wages and the productivity shock (the last argument of equation A8). Hence, labor pooling generates, as long as shocks are not perfectly correlated across firms, a pecuniary externality that increases expected profits. Note that risk (specific to the firm) is the crucial element behind this externality that will play the central role in the tendency for agglomeration.

The Location Equilibrium

Equations A6 and A8 give the expected wage and profit in each location. For a location configuration to be an equilibrium, both mobile firms and mobile workers must have no incentive to move. We assume that both must decide where to locate before shocks occur, which implies that expected wages and expected profits must be equalized across locations. We define $\lambda_L = L_A/L$ as the share of workers in location A and, similarly, $\lambda_F = n_A/N$ as the share of firms in location A. For expected wages to be identical, it must be that $\lambda_L = \lambda_F$, so that the share of firms is equal to the share of workers. The reason is that if firms move to one location, the wage rate will increase in that location, attracting workers up to the point where wages are equalized. This is the first equilibrium relation, which we call the *WW* schedule.

The equilibrium relation that implies that expected profits are equalized is a bit more complex. Solving $E\pi_A = E\pi_B$, we obtain a second relation between λ_L and λ_F, which we call the $\pi\pi$ schedule:

$$(A9) \qquad (1-2\lambda_F)[\gamma^2L^2\lambda_F^2 - N1 - \lambda_F\lambda_F\sigma_\varepsilon^2 + \gamma\mu F^2]$$
$$-\gamma^2L^2\lambda_F^2(1-2\lambda_L) = 0.$$

It can be shown that, as long as it is continuous (that is, σ_ε^2 is not too high), the $\pi\pi$ schedule defines a positive relation between λ_L and λ_F. If workers move to a location, they will decrease wages in that location so that firms will also move. However, this positive relation between the share of workers and the share of firms in region A is nonlinear because of the presence of firm-specific shocks (the labor market pooling effect) and the impact on the land market. If these two effects were absent ($\sigma_\varepsilon^2 = \mu = 0$), the location equilibrium would then be indeterminate.

An obvious equilibrium exists—the symmetric equilibrium ($\lambda_L = \lambda_F = 1/2$) or the dispersed equilibrium. As usual in this literature, we can ask under which conditions this equilibrium is stable and under which conditions other equilibria exist, with one location arising as the main agglomeration with a majority of workers and firms.

Let us first consider the case where the symmetric equilibrium is stable. Estimated at the symmetric equilibrium, the slope of the $\pi\pi$ schedule is:

$$(A10) \qquad \frac{\partial\lambda_F}{\partial\lambda_L} = \frac{\gamma^2L^2}{\gamma^2L^2 + 4\gamma\mu F^2 - N\sigma_\varepsilon^2}.$$

Suppose first that the slope of the $\pi\pi$ schedule is less than the slope of the *WW* schedule. If we assume that workers move to the location with the higher expected wage, and that firms move to the location that offers higher expected profits, then the dynamics of the model are shown by the arrows in Exhibit 1.

The symmetric equilibrium is stable in the sense that if firms change locations, profits will increase in the location from which they leave and decrease in the other location. The same would happen to wages if workers were to leave a location. The slope in equation A10 is less than 1 when:

$$(A11) \qquad 4\gamma\mu F^2 > N\sigma_\varepsilon^2.$$

Not surprisingly, no agglomeration process takes place when: 1) wages and rents in a location react strongly to an increase in demand by firms relocating to that location

(high γ, μ, and F), and 2) when N and σ_ε^2 are so small that the benefits of labor pooling for firms are small either because there are few firms to pool with or because the specific firms' shocks are small. It can be shown that the inequality in equation A11 is the same one that implies that no other equilibrium exists. This scenario is depicted in Exhibit 2.

When the inequality of equation A11 is reversed, the $\pi\pi$ schedule becomes steeper than the WW curve at the symmetric equilibrium and two other interior asymmetric equilibria appear, which are given by:

$$\text{(A12)} \qquad \lambda_F = \lambda_L = \frac{1}{2} \pm \frac{\sqrt{N^2\sigma_\varepsilon^4 - 4\gamma\mu NF\sigma_\varepsilon^2}}{2N\sigma_\varepsilon^2}.$$

Hence, one of the locations becomes an agglomeration, which attracts most workers and most firms. The equilibrium is such that wages and profits are the same in both locations. From equation A5, one can see that the price of land is higher in location A than in location B. The difference in land price reflects the benefit from labor pooling in location A.

The dynamics of the model that exist when agglomeration forces are strong enough are depicted in Exhibit 3. It can be shown that the two asymmetric equilibria are stable. As is usual in models of new economic geography with multiple equilibria, the identity of the location that ends up being the agglomeration is indeterminate. It can be interpreted loosely as the result of history and/or an initial small advantage. We choose location A as the location where most firms and workers concentrate.

Note that the higher the risk specific to firms (the higher σ_ε^2), the more firms and workers will want to concentrate in one location only. Hence, increased uncertainty is not, in itself, a force that leads firms to leave the urban, densely populated areas. In fact, this simple model shows that because of market pooling arguments, it may lead to an agglomeration process making the symmetric equilibrium unstable, and it may increase the attractiveness of the concentrated area once the agglomeration process is achieved.

To consider the possible effect of terrorism on the structure of cities, we start from the assumption that such an agglomeration phenomenon has indeed occurred, exemplified by New York City, so that location A has most firms and workers and that this equilibrium is initially stable.

There are different ways to interpret the impact of the events of September 11 on cities in terms of our model: 1) as an unexpected, negative onetime shock to firms based in the concentrated area, 2) as a permanent increase in expected costs in location A, the largest city, as future terrorist attacks might target this location (for example, because of the presence of financial markets, government activities, or other "symbolic activities"), 3) as a permanent increase in the disutility of workers who leave the concentrated location A, and 4) as a permanent increase in costs to firms as a function to the number of other firms in the same location.

An Unexpected, Negative Onetime Shock

In the first interpretation, and if we take strictly our assumption that the location choice is made before the shocks are realized, a onetime negative shock to profits in location A has no impact on location. Firms' profits decrease in location A. If the shock also affects labor productivity temporarily, wages in location A will decrease. A onetime unexpected shock has an impact on profits and wages, but not on the location equilibrium that results from a long-term choice. Another way to say this is that a temporary shock does not affect the economic forces that explain the agglomeration process in the first place (that is, the parameters that explain the instability of the symmetric dispersed equilibrium), nor does it affect the extent of spatial concentration (that is, the parameters that affect the share of firms and workers in location A in equation A12).

A Permanent Increase in Costs in Large Cities

In the second interpretation, we assume that costs increase permanently in location A by an amount T_A. This could be due to the fact that the threat of terrorism induces higher costs in terms of security and insurance. Using equation A9, we see that the $\pi\pi$ schedule then becomes:

$$\text{(A13)} \qquad (1-2\lambda_F)[\gamma^2L^2\lambda_L^2 - N(1-\lambda_F)\lambda_F\sigma_\varepsilon^2 + \gamma\mu F^2] - \gamma^2L^2\lambda_F^2(1-2\lambda_L) - 2T_A\gamma N^2\lambda_F^2(1-\lambda_F)^2 = 0.$$

It can be shown that the $\pi\pi$ schedule is shifted downward and becomes flatter, and that $\partial\lambda_F/\partial T_A$ is always negative when evaluated at the "high" equilibrium, that is, in the equilibrium with concentration in location A. This is shown in Exhibit 4.

In this case, the equilibrium becomes less concentrated as firms leave location A for location B. However, agglomeration forces are still present, and location A remains the largest city despite the fact that due to the terrorism threat, the costs of firms have increased permanently. The fact that agglomeration gains (the labor pooling argument) still exist implies that one location is bound to have most firms and workers. In models of economic geography with multiple equilibria, history provides a strong advantage to the city with the initial concentration. This model shares this property. Another way to say this is that an agglomeration rent exists in the city: even if in the specific city the profits of firms are hit by a permanent shock to costs, most firms will want to stay to benefit from labor pooling. Of course, if the shock is large enough, that is, if the $\pi\pi$ schedule moves downward sufficiently, the equilibrium of concentration in location A could disappear. In this case, agglomeration would shift to the other possible equilibrium, location B. But the agglomeration gain derived from labor pooling would still imply an equilibrium with high concentration: another way to say this is that the symmetric equilibrium remains unstable.

Note that profits also decrease in location B when location A experiences this permanent decrease in profits. To show how the adjustment takes place, we write the level of profits in A (we know that in equilibrium profits must be equal in B) as a function of the share of workers and firms in region A:

$$\text{(A14)} \qquad \pi_A = \frac{\gamma}{2}\frac{\lambda_L^2}{\lambda_F^2}\frac{N^2}{L^2} + \left(1 - \frac{1}{N\lambda_F}\right)\frac{\sigma_\varepsilon^2}{2\gamma} + F^2\mu\gamma\frac{N^2}{\lambda_F^2} - T_A - tN\,.$$

The first element of the profit level is the one linked to labor use by the firms. In equilibrium, we know that $\lambda_F = \lambda_L$, so that the permanent increase in costs through T_A does not alter profits in an indirect way through labor use and wages. The reason is that as firms leave location A, workers also leave, so that wages are unaffected. The second element is linked to the agglomeration benefits that come from labor pooling. As λ_F decreases due to the increase in T_A, these agglomeration benefits also decrease, which puts further downward pressure on profits in location A. The third element is linked to the cost of land. As firms leave location A, that is, as λ_F decreases, the demand for land decreases as well as its cost. Hence, this mitigates the fall in profits in A.

In location B, for the same symmetric reasons, wages are not affected, the agglomeration benefits of labor pooling increase, and land costs increase. In equilibrium, the cost-of-land increase in location B must be large enough to more than compensate for the positive effect of the labor pooling mechanism on profits.

To infer the transitional dynamics implied by the permanent increase in costs in location A, we retain the assumption that firms move with profit changes and that workers move with wage changes. In this case, when costs to firms rise, profits decrease and the equilibrium goes from 1 to 2 (Exhibit 3), so that firms leave the large city. This decreases wages in the city and leads workers to follow the firms at point 3, which can be considered the long-term equilibrium. Note that there is an amplifying effect due to the interaction of firms and workers. More firms leave in the long term than in the short term (λ_F is lower at point 3 than at point 2): the reason is that as workers leave city A, labor supply goes down and wages increase. This leads more firms to move as labor costs increase in city A.

Note that in this interpretation, wages are not affected in the long-term equilibrium but decrease in location A in the short term. If firms move first and workers second, wages will overshoot downward in the short term in location A and then increase back to the initial level.

The fact that wages do not change in the long term may not be true if some of the labor pooling mechanism were to increase labor productivity and therefore the demand for workers. In this case, wages would be negatively affected in the long term by the fact that firms and therefore workers become less spatially concentrated.

The numerical example described in this article uses these parameters: $L = 10$, $N = 5$, $\sigma_\varepsilon^2 = 0.5$, $\mu = 2$, $\gamma = 0.4$, $F = 0.7$, $\beta = 1.3$, $\delta = 1.5$, $c = 0.19$. With these parameters, both the average wage and productivity and its variance at the firm level are 0.5.

A Permanent Increase in Perceived Risk to Agents in Large Cities

Up to now, we have assumed that agents, when choosing where to live, only care about wages. Obviously, this does not take into account other important aspects of location choice. The concern for safety should be an important one, and one way to interpret the events of September 11 is that the perception of safety has decreased permanently in large cities. Suppose we now assume that agents equalize utility rather than wages, and

that utility is simply the wage minus a risk parameter (call it R_A), which was permanently increased in New York City after the attack. City A becomes a perceived target because it is the biggest city but also presumably because of other specific factors such as the presence of financial markets. Hence, the *WW* schedule is now changed. It becomes:

$$\text{(A15)} \qquad \lambda_F^2 N R_A + \lambda_F(\gamma L - N R_A) - \gamma L \lambda_L = 0 .$$

There is still a positive relation between λ_F and λ_L, which is given in Exhibit 5. As the *WW* schedule shifts up from "zero risk" to the schedule with the specific risk associated with city A perceived by agents, the equilibrium is such that both workers and firms move out of the city. The equilibrium shifts from point 1 to 2 in Exhibit 5. The dynamics are such that as workers move out of the large city, wages increase there with a negative effect on profits, so that firms also move out. Note, however, that in equilibrium, the share of firms in the big city is larger than the share of workers, so that wages are higher in city A than in city B. The difference in wages is the risk factor. The reason for this is that to retain workers in city A, firms must pay higher wages there. However, utility decreases in both locations.

The increase in wages has a negative effect on profits. It can be shown that firms' profits decrease in both locations. This comes from higher wages in the big city and also from the fact that the gains from labor pooling have declined with lower concentrations of workers and firms in the big city. From that point of view, economic geography becomes less efficient, as it was in the previous case.

Again, a large enough shock could make the big city "disappear" and reappear in location B. If we exclude this unlikely scenario, the important conclusion to draw is that because the threat of terrorism does not eliminate the labor pooling gains, the symmetric equilibrium with dispersion will not emerge as a stable equilibrium and city A will remain the location with most firms and workers.

An Increase in Costs Linked to Industrial and Population Density

In the previous interpretations of the economic effects of September 11, the increased risk factor either to firms or agents was specific to a city, city A in our example. We now analyze how the results differ when the effect on firms' costs is linked not to a specific city but to the density of the location, based on the idea that cities with more firms are more likely to be the target of terrorist attempts. Hence, we now assume that profits are given by:

$$\text{(A16)} \qquad \pi_{iA} = R_{iA} - w_A L_{iA} - T(2\lambda_F - 1)$$

$$\pi_{iB} = R_{iB} - w_B L_{iB} - T(1 - 2\lambda_F)$$

$$T(0) = 0 ; \; T' > 0 .$$

The condition that expected profits and expected wages are equal across locations still determines the location equilibrium. The symmetric equilibrium with $\lambda_F = \lambda_L = 1/2$ is a possible equilibrium. The $\pi\pi$ schedule has the same form as before but is flatter, as its slope, when estimated in the symmetric equilibrium $\lambda_F = \lambda_L = 1/2$, is now:

$$\text{(A17)} \qquad \frac{\partial \lambda_F}{\partial \lambda_L} = \frac{\gamma^2 L^2}{\gamma^2 L^2 + 4\gamma\mu F^2 - N\sigma_\varepsilon^2 + \gamma L^2 T'(2\lambda_F - 1)} .$$

The fact that locations with a higher share of firms are more likely to be targeted and therefore see the costs of doing business increase is similar to a congestion cost.[16] This interpretation is different from the one in the previous section, where the terrorism threat did not alter the fundamental benefit of being spatially concentrated. In this interpretation, the two equilibria with spatial agglomeration could disappear and the dispersed equilibrium could become the unique stable equilibrium if the expression in equation A17 becomes less than 1. In this case, Exhibit 2 would become the valid description of the equilibrium.

This interpretation of September 11 basically says that the benefits of labor pooling (and others not modeled here) would be overcome by the congestion cost linked to increased security, itself linked to the share of firms in the city. In this interpretation, a small negative shock to the large city would make firms leave. This would depress wages there and lead workers to move out. As firms leave the city, the gains from labor pooling unravel and the dispersed equilibrium with two cities of equal size becomes the unique stable equilibrium. For such a scenario to become possible, the cost linked to the terrorist threat in a city needs to increase strongly with the share of firms in that city: $T'(2\lambda_F - 1)$ must be high enough. In this case, spatial dispersion becomes an insurance device.

A close analysis of the preceding scenario would interpret the threat of terrorism as a decrease in the variance of firms' specific shocks relative to shocks specific to a location, which affects all firms. In this case, the very benefits of spatial agglomeration (at least those of labor pooling) would be reversed: the insurance mechanism that led firms to agglomerate would lead them to disperse.

In the scenario in which the increase in costs to large cities is not enough to make the symmetric equilibrium a stable equilibrium, the analysis is similar to the one in this article. The large city loses firms and workers to the smaller city but remains the location with the majority of firms and workers. The analysis would be similar in the case where the cost affects workers rather than firms.

The Core Periphery Model

Our second model, the core periphery model of Fujita, Krugman, and Venables (1999), offers one explanation for agglomeration. The ingredients of the model are two locations, each endowed with an equal amount of an immobile factor (the authors call this factor agricultural labor). There are two goods in the model: a numeraire, produced under constant returns and traded freely (the agricultural good), and a differentiated manufactured good, produced under increasing returns with monopolistic competition modeled à la Dixit-Stiglitz. Unlike the labor pooling model, the core periphery model is a general equilibrium model, where all agents maximize and all markets clear.

Shipping costs for manufactured goods are modeled as an iceberg. This means that to deliver one unit to the other region, a manufacturer must ship $T > 1$ units, $T - 1$ of which "melt" in transit. In addition, firms in the big city must pay an iceberg "terror tax," $\tau > 1$: to produce one unit of a good that can be sold, each firm must spend $\tau - 1$ units of output on security, insurance, and associated inefficiencies.

Introducing some necessary notation, we give the statement of the model by way of a system of eight equations:

$$\text{(A18)} \quad Y_A = \mu\lambda w_A + \frac{1-\mu}{2}$$

$$\text{(A19)} \quad Y_B = \mu(1-\lambda) w_B + \frac{1-\mu}{2}$$

$$\text{(A20)} \quad G_A^{1-\sigma} = \lambda(\tau w_A)^{1-\sigma} + (1-\lambda)(T w_B)^{1-\sigma}$$

$$\text{(A21)} \quad G_B^{1-\sigma} = \lambda(T\tau w_A)^{1-\sigma} + (1-\lambda) w_B^{1-\sigma}$$

$$\text{(A22)} \quad w_A^{\sigma} = Y_A\left(\frac{G_A}{\tau}\right)^{\sigma-1} + Y_B\left(\frac{G_B}{T}\right)^{\sigma-1}$$

$$\text{(A23)} \quad w_B^{\sigma} = Y_A\left(\frac{G_A}{T\tau}\right)^{\sigma-1} + Y_B G_B^{\sigma-1}$$

$$\text{(A24)} \quad \omega_A = \frac{w_A}{G_A^{\mu}}$$

$$\text{(A25)} \quad \omega_B = \frac{w_B}{G_B^{\mu}}.$$

This is less imposing than it looks at first, although the model can only be solved numerically. Equations A18 and A19 give income in each region, which is composed of labor income from manufactured workers plus the wages of agricultural workers. The total number of manufacturers, and the share of manufactured goods in Cobb-Douglas utility, is μ. The share of mobile workers located in region A is λ, and w_r is the nominal wage in region $r \in \{A, B\}$. The number of agricultural workers in each region is 1/2.

The next two equations define the price indexes for manufactured goods in each region, G_r, and are derived from a CES (constant elasticity of substitution) utility function with parameter $\sigma > 1$ and the assumption of iceberg transport costs.

The nominal wage equations A22 and A23 are a consequence of the zero-profit equations for manufactured goods. Finally, equations A24 and A25 give the definition of real wages, ω_r, in each location; recall that agricultural goods have a price of 1 and a weight of $(1-\mu)$ in the utility function of workers.

The equilibrium division of labor across regions is what the model aims to explain. Since the manufactured good is costly to ship between regions, firms will want to locate near their customers. If one region has more firms than the other, then the cost of living will be lower there because fewer goods incur transport costs, which makes it an attractive place for workers to locate. These two forces for agglomeration are called the "market access" and "price index" effects, respectively. They are balanced by a force that favors dispersion—that is, firms prefer less competition, which gives them an incentive to locate away from other firms.

We model terrorism in the city as an increase in the cost of doing business there. To see how this affects equilibrium agglomeration, start in the core periphery equilibrium and consider an increase in τ from its initial level of 1. The model is analytically intractable, but we can illustrate numerically the result that you would expect: a small increase reduces the real wage gap between the city and hinterlands but does not overturn it, while a larger increase reverses the real wage gap and causes the city to cease to exist.

In particular, consider starting at the core periphery equilibrium with $T = 1.7$, $\sigma = 5$, and $\mu = 0.4$. With $\tau = 1$ (the September 10 equilibrium), workers in location A get a real wage premium of 0.02. If we increase the tax to $\tau = 1.05$, the real wage premium falls to 0.004, and if we increase the tax to $\tau = 1.10$, the real wage gap becomes -0.012. So a terror tax of 5 percent has no effect on agglomeration, while a 10 percent tax causes location A to spiral downward, which only stops when $\lambda = 1/2$. With these parameters, $\lambda = 1/2$ is a stable equilibrium, so even if the terror tax is removed, location A will remain no larger than B if the terror tax was onerous enough to cause a disagglomeration.

Note that it is not terrorism per se that causes people to leave the city; rather, it is the loss of the economic rationale for agglomeration that causes workers and firms to disperse. In the presence of a terror tax, the benefits of agglomeration (market access, cost of living) are counteracted: the tax reduces the effective size of the market that firms can access and raises the cost of living for workers. Note, though, that the cost of living also rises for workers in the hinterlands, since they must now pay the terror tax on goods that they import (along with the usual transport costs). So, for a given distribution of firms, the relative cost of living does not change, but the smaller effective market size in the city reduces the nominal wage gap between city and hinterland labor.

Endnotes

1. Although distance does not matter for some financial transactions, it still matters for others. For example, Petersen and Rajan (2000) show that most bank lending still occurs over fairly short distances, although the average distance between lender and borrower has increased over time.

2. Gaspar and Glaeser (1998) offer another reason why there may be agglomeration of producers of weightless goods, such as financial services and software. They argue that information technology may be a complement to, rather than a substitute for, face-to-face interaction, so that greater ease of electronic communication *increases* the incentive to locate near other producers.

3. Although we concentrate on the labor pooling and increasing returns/transport costs motives for agglomeration, we acknowledge that there are other economic explanations for city formation. The most well-known is the spillover model, set out most thoroughly in Henderson (1988). In Henderson's model, agglomeration raises the productivity of firms that locate near other firms in the same line of business.

4. This is purely a simplifying assumption. If workers also disliked uncertainty, our results would actually be stronger. We return to the issue of workers' risk-aversion in our discussion of the effects of terrorism.

5. There is also a mirror-image equilibrium where $\lambda_F = \lambda_L < 1/2$, but we omit it to simplify the analysis.

6. We do not assume that the smaller city is free from the threat of terrorism, nor do we assume that the terror tax falls to zero once firms cross the city limits. Our analysis simply supposes that the tax is higher in big cities than it is elsewhere.

7. From Haughwout and Inman (forthcoming), it is possible to infer that a lower bound for this ratio is around 7. For New York City, it is much higher. The results are not very sensitive to this inference.

8. Profits decrease by 0.2 percent in both locations, as does economywide GDP. The impact on profits and GDP is less than the direct effect on costs because firms can and do take advantage of the alternative location where they can avoid the terror tax.

9. More precisely, we assume in this case that workers equalize utility across locations. Utility is the wage minus the fear factor.

10. The fear factor should not be interpreted here as the loss due to terrorism felt by all in the country. In a much narrower sense, it is the difference of utility due to a future terrorism threat perceived by those living in the city.

11. The wage increases in the city by 0.2 percent and decreases in the nonurban area by 0.4 percent. Profits decrease in both locations by 0.8 percent. In the city, this occurs because of higher wages and lower agglomeration gains; in the nonurban areas, it occurs because of higher rents. The ratio of rents between the two locations decreases from 10 to 7.4. Again, this is a stabilizing factor for the city.

12. Regions A and B in this model differ from those in the labor pooling model.

13. Residents of large cities, particularly New York, may disagree with the notion that the cost of living is lower in these cities. Here, we abstract from land and housing prices, and interpret transport costs broadly to include the costs of consuming a wide variety of goods. For goods that are available only in the city (such as fine restaurants and unique entertainment), their cost for those outside the city includes airfare and hotel bills.

14. There is one other equilibrium between $\lambda_L = 0.5$ and $\lambda_L = 1$ where the real wage gap is zero, but it is unstable.

15. This is the interpretation of the post-war Japanese experience offered by Davis and Weinstein (forthcoming).

16. If we were to make this cost a function of the share of workers in each location, the qualitative results would be similar.

References

Bram, Jason, Andrew Haughwout, and James Orr. 2002. "Has September 11 Affected New York City's Growth Potential?" Federal Reserve Bank of New York Economic Policy Review 8, no. 2 (November): 81-96.

Bram, Jason, James Orr, and Carol Rapaport. 2002. "Measuring the Effects of the September 11 Attack on New York City." Federal Reserve Bank of New York Economic Policy Review 8, no. 2 (November): 5-20.

Davis, Donald R., and David E. Weinstein. Forthcoming. "Bones, Bombs, and Breakpoints: The Geography of Economic Activity." American Economic Review.

DRI-WEFA. 2002. "The Role of Metro Areas in the U.S. Economy." Report prepared for the U.S. Conference of Mayors.

Fujita, Masahisa, Paul Krugman, and Anthony J. Venables. 1999. The Spatial Economy: Cities, Regions, and International Trade. Cambridge: MIT Press.

Gaspar, Jess, and Edward L. Glaeser. 1998. "Information Technology and the Future of Cities." Journal of Urban Economics 43, no. 1 (January): 136-56.

Glaeser, Edward L., and Jesse M. Shapiro. 2002. "Cities and Warfare: The Impact of Terrorism on Urban Form." Journal of Urban Economics 51, no. 2 (March): 205-24.

Haughwout, Andrew, and Robert Inman. Forthcoming. "Should Suburbs Help Their Central City?" Brookings-Wharton Papers on Urban Affairs.

Henderson, J. Vernon. 1988. Urban Development. Oxford: Oxford University Press.

Hobijn, Bart. 2002. "What Will Homeland Security Cost?" Federal Reserve Bank of New York Economic Policy Review 8, no. 2 (November): 21-33.

Krugman, Paul. 1991. Geography and Trade. Cambridge: MIT Press.

Petersen, Mitchell A., and Raghuram G. Rajan. 2000. "Does Distance Still Matter? The Information Revolution in Small Business Lending." NBER Working Paper no. 7685, May.

Standard and Poor's Insurance. 2002. "U.S. Property/Casualty Insurance Midyear Outlook 2002: Negative Fundamentals Outweigh Higher Pricing." May.

THE FEDERAL RESERVE BANK OF NEW YORK
ECONOMIC POLICY REVIEW

Single-copy subscriptions to the *Economic Policy Review* (ISSN 0147-6580) are free. Multiple copies are available for an annual cost of $12 for each additional subscription. Checks should be made payable in U.S. dollars to the Federal Reserve Bank of New York and sent to the Public Information Department, 33 Liberty Street, New York, NY 10045-0001 (212-720-6134). Single and multiple copies for U.S. subscribers are sent via third- and fourth-class mail.

Library of Congress Card Number: 77-646559