Rajashri Chakrabarti

Program Design, Incentives, and Response: Evidence from Educational Interventions

- In an effort to reform K-12 education, policymakers have introduced vouchers in some U.S. school districts, enabling students to transfer from public to private schools.
- The different designs of two school voucher programs—the Milwaukee and Florida programs—have had different effects on public school incentives and performance.
- In Milwaukee, vouchers were imposed from the outset; in Florida, schools were first threatened with vouchers and thus had an incentive to avoid them.
- The Florida public schools' efforts to avoid vouchers resulted in performance effects that far exceeded those of Milwaukee's program.
- Program design is critical: Policies that present failing public schools with functional and credible sanctions are best suited to provide the results intended by policymakers.

1. Introduction

Concerns that U.S. students are not performing as well as their counterparts in other developed countries on international math and science tests have led to widespread demands for the reform of K-12 education in the United States. Of the various reforms under consideration, school voucher reform is at the forefront.

Vouchers are scholarships that make students eligible to transfer from public to private schools. A basic feature of all publicly funded voucher programs in the United States is the funding of vouchers by public school revenue, so that money always "follows" students. In other words, schools that lose students lose their corresponding funding. Schools therefore recognize the financial implications of vouchers and have an incentive to avoid being subject to voucher programs.

This article investigates the role of program design in the context of two such educational interventions in the United States—the Milwaukee and Florida school voucher programs—and analyzes the effects of design on public school incentives and performance.[1] We demonstrate that variations in program design have markedly different outcomes for public schools affected by vouchers.

The Milwaukee program, introduced in 1990, was the first voucher program in the country. Implemented in 1999, the

Rajashri Chakrabarti is an economist at the Federal Reserve Bank of New York.
rajashri.chakrabarti@ny.frb.org

The author thanks Andrew Haughwout for helpful suggestions. Noah Schwartz provided excellent research assistance. All errors are the author's. 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.

Florida program was the nation's third, following Cleveland's. The Milwaukee and Florida voucher programs share the basic feature of funding by public school revenue. But there are crucial differences. Milwaukee's is a means-tested program targeting low-income students while Florida's embeds a voucher program in a full-fledged accountability system.

Using test-score data from Milwaukee and Florida and implementing a difference-in-differences estimation strategy, our study estimates the impact of each program by comparing the post-program results of the affected schools with a comparable set of control schools. Controlling for potentially confounding pre-program time trends and post-program common shocks, we find that the performance effects of the Florida program far exceed those of Milwaukee's program. These results are quite robust in that they hold after controlling for other confounding factors, such as mean reversion and a possible stigma effect; they also withstand several sensitivity tests.

Our findings have important policy implications, which we consider in the context of New York State's federal, state, and city accountability programs. These programs include New York City's accountability policy, known as the "Progress Report" policy, and the federal No Child Left Behind (NCLB) law, as implemented by New York State.

Our study is organized as follows. Section 2 describes the Milwaukee and Florida voucher programs. In Section 3, we discuss the incentives created by the programs and the corresponding responses that might be expected from the affected public schools. Our data and empirical strategy are reviewed in Sections 4 and 5, respectively. Section 6 presents our results, and Section 7 considers policy implications.

2. Institutional Details

The first publicly funded school voucher program in the United States was established in Milwaukee, Wisconsin, in 1990. The Milwaukee Parental Choice Program made the city's entire low-income public school population eligible for vouchers. Specifically, starting in the 1990-91 school year, the program made all Milwaukee public school students with family income at or below 175 percent of the poverty line eligible for vouchers to attend nonsectarian private schools.

In contrast, the Florida Opportunity Scholarship Program, introduced in 1999, can be looked upon as a "threat-of-voucher" program. Here, failing public schools were threatened with the imposition of vouchers, with vouchers implemented *only if* schools failed to meet a government-designated cutoff quality level. The institutional details of the Milwaukee and Florida programs are summarized in Table 1.

The Florida Department of Education classified schools according to five grades: A, B, C, D, or F. The state assigned school grades based on Florida Comprehensive Assessment Test (FCAT) reading, math, and writing scores. For FCAT reading and math, it categorized students into five achievement levels—1 lowest, 5 highest—that correspond to specific ranges on the raw-score scale. Using current-year data, the Department of Education assigned an "F" grade to a school if it was below the minimum criteria in reading, math, and writing; a "D" if it was below the minimum criteria in one or two of the three subject areas; and a "C" if it was above the minimum criteria in all three subjects, but below the higher performing criteria in all three. In reading and math, at least 60 percent (50 percent) of students had to score level 2 (3) and above; in writing, at least 50 percent (67 percent) had to score 3 and above to meet the minimum (high-performing) criteria in that respective subject.[2]

Under the Florida Opportunity Scholarship Program, all public school students became eligible for vouchers, or

The major difference in program design between the Milwaukee and Florida programs is that in Milwaukee vouchers were imposed at the outset, whereas in Florida failing schools were first threatened with vouchers, with vouchers introduced only *if the schools failed to show adequate improvement in performance.*

[1] Our study focuses on the impact of alternative voucher designs on public school performance. A growing body of literature analyzes the many issues associated with school vouchers. Nechyba (1996, 1999, 2000, 2003) analyzes distributional effects of alternative voucher policies in a general equilibrium framework; Epple and Romano (1998, 2002) and Chakrabarti (2009) investigate sorting attributable to vouchers; Manski (1992) considers the impact of vouchers on public school expenditure and social mobility; and McMillan (2004) and Chakrabarti (2008b) model the quality of public schools facing vouchers. Hoxby (2003a, b) and Chakrabarti (2008a) study the effects of the Milwaukee voucher program, while Greene (2001), Greene and Winters (2003), Figlio and Rouse (2006), West and Peterson (2005), and Chakrabarti (2007, 2008a) study the effects of the Florida program.

[2] In 1999, seventy-eight schools received an "F" grade. Students in two of those schools became eligible for vouchers. In 2000, four elementary schools received an "F," although none became eligible for vouchers. In 2001, no schools received an "F" grade. In 2002, sixty-four schools received an "F." Students in ten of those schools became eligible for vouchers. In 2003, students in nine schools became eligible for vouchers; in 2004, the figure was twenty-one schools.

Table 1
Comparison of Milwaukee and Florida Voucher Programs

Milwaukee Program	Florida Program
• First U.S. voucher program • Started in 1990-91 school year • Public school students with family income at or below 175 percent of the poverty line eligible for vouchers to attend nonsectarian private schools • Private schools were not permitted, by law, to discriminate against students who apply with vouchers: – Had to accept all students unless oversubscribed – If oversubscribed, had to choose students randomly • Average voucher amount equaled the state aid per pupil, and vouchers were financed by an equivalent reduction of state aid to the school district • 1990-91 and 1996-97: – Average voucher amounts were $3,346 – Vouchers as a percentage of total revenue per pupil were 45.23 percent	• Third U.S. voucher program • Started in 1998-99 school year • Vouchers contingent on school performance • Schools classified according to five grades: A, B, C, D, F (A-highest, F-lowest) – Grades based on the Florida Comprehensive Assessment Test (FCAT) reading, math, and writing scores – F, if below the minimum criteria in reading, math, and writing – D, if below the criteria in one or two of the three subjects – C, if above the minimum criteria in all three subjects, but below the higher performing criteria in all three • Students categorized into five achievement levels in FCAT reading and math (1-lowest, 5-highest) • Minimum criteria: – Reading and math: at least 60 percent must score level 2 and above – Writing: at least 50 percent must score level 3 and above • High-performing criteria: – Reading and math: at least 50 percent must score level 3 and above – Writing: at least 67 percent must score level 3 and above • All students of a public school became eligible for vouchers if the school received two "F" grades in a period of four years • Private schools were not permitted, by law, to discriminate against students who apply with vouchers – Had to accept all students unless oversubscribed – If oversubscribed, had to choose students randomly • Average voucher amount equaled the state aid per pupil, and vouchers were financed by an equivalent reduction of state aid to the school district • 1999-2000 and 2001-02: – Average voucher amounts were $3,330 – Vouchers as a percentage of total revenue per pupil were 41.55 percent

Source: Information and data provided in various Florida Department of Education and Milwaukee Department of Public Instruction reports.

"opportunity scholarships," if the school received two "F" grades in a period of four years. Therefore, a school that received an "F" for the first time was exposed to the threat of vouchers, but did not face them unless and until it got a second "F" within the next three years. Thus, the major difference in program design between the Milwaukee and Florida programs is that in Milwaukee vouchers were imposed at the outset, whereas in Florida failing schools were first threatened with vouchers, with vouchers introduced *only* if the schools failed to show adequate improvement in performance.

Apart from the above differences, the design of the two programs was strikingly similar. In both programs, private schools could not, by law, discriminate against students who applied with vouchers—the schools had to accept all students unless they were oversubscribed, in which case they had to choose students randomly. Indeed, the application form did not ask questions about the student's race, sex, parents' education, past scores, or prior records (for example, truancy, violence). The questions were specifically worded only to ascertain whether the student was eligible for the program.[3] The system of funding for the Milwaukee and Florida voucher programs was also very similar. Under each program, the average voucher amount was equal to the state aid per pupil, and vouchers were financed by an equivalent reduction of state aid to the school district. Thus, state funding was directly tied to student enrollment, and enrollment losses due to vouchers were reflected in a revenue loss for the public school.[4] The average voucher amounts under the Milwaukee (1990-91 through 1996-97) and Florida (1999-2000 through 2001-02) programs were $3,346 and $3,330, respectively. During these periods, vouchers as a percentage of total revenue per pupil were 45.23 percent in Milwaukee and 41.55 percent in Florida.

[3] While the schools could not employ any selection criteria for the voucher students, this was not the case for nonvoucher students in the same school. Also note that the private schools had the choice of whether to participate in the program. However, if they decided to participate, they were required by law to accept all students or to choose students randomly, if oversubscribed.

3. Discussion: Effects of the Programs on Public School Incentives and Response

What incentives would be created by the aforementioned program rules, and how would one expect the affected public schools to respond? Consider a public school subject to the Florida program, a school that has just received its first “F” grade (“F-school” hereafter). The school realizes that if it can avoid another “F” grade in the next three years, it can escape vouchers and the monetary loss and embarrassment associated with them.[5] Therefore, it would have an incentive to improve its scores so as not to receive a second “F” grade. In contrast, if the same school were subject to a Milwaukee-type voucher program—in which vouchers have already been introduced—the school could not avoid vouchers (and the revenue loss)

> *In a Florida-type [voucher] program, the threatened public schools . . . have more of an incentive to respond in order to improve their scores and escape vouchers.*

altogether. In this case, improvements would serve to retain or attract a few students, but the effect would be marginal compared with that of a Florida-type program. In a Florida-type program, the threatened public schools (schools that have received their first “F” grade) have more of an incentive to respond in order to improve their scores and escape vouchers.[6] Thus, the key difference between the two programs is that in the Milwaukee program, vouchers have already been implemented, whereas the Florida program first threatens the schools and gives them a window to respond, and an adequate response can preclude sanctions. Sanctions (vouchers) are implemented only if the schools fail to attain the predesignated standard.

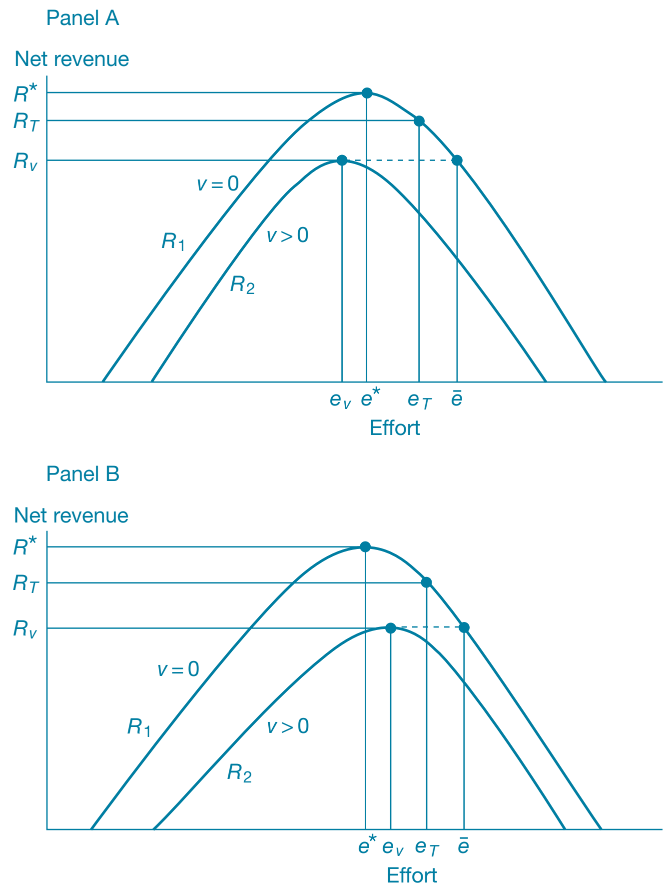
[4] We focus on the Milwaukee program up to 1996-97. The reason is that following a 1998 Wisconsin Supreme Court ruling, there was a major shift in the program when religious private schools were allowed to participate. Moreover, the financing of the Milwaukee program underwent some crucial changes, so that the voucher amounts and the revenue loss per student due to vouchers were not comparable between the Florida and second-phase Milwaukee programs. See Chakrabarti (2008a) for an analysis of how the shift in the Milwaukee voucher program affected public school incentives and performance as well as for a comparison of public school responses in the two phases of the program. We focus on the Florida program up to 2001-02. This period is chosen because schools that received an “F” grade in 1999 would face the threat of vouchers only through 2002.

[5] The loss of students due to vouchers leads to a decrease in both revenue and costs for the school. But for a school operating at full capacity, the cost savings due to the loss of students are marginal, while the loss in revenue is not. This effect is a major reason why public schools do not support vouchers.

[6] For a formal proof, see Chakrabarti (2008b).

Chart 1

Analyzing the Effect of “Voucher Threat” versus Vouchers



The intuition above is shown in Chart 1. Let R_1 illustrate the initial net revenue function of the public school. The public school chooses the effort to maximize net revenue. Let this equilibrium effort be denoted by e^* and the corresponding net revenue by R^*. Now assume that Milwaukee-type vouchers are introduced. This leads to a downward shift of the net revenue function—the new net revenue function is denoted by R_2 and the corresponding optimum effort and net revenue by e_V and R_V, respectively.[7] Panel A of the chart illustrates the case in which $e_V < e^*$, and panel B the case in which $e_V > e^*$. The chart implies that any target effort in the range $(e_V, \bar{e}]$ under a threat-of-voucher regime will induce an effort strictly greater than e_V. For example, assume that the policymaker implements a target effort, e_T. Satisfying this target would lead to a net revenue of R_T while failing to satisfy it would lead to the introduction of vouchers and corresponding revenue of $R_V (< R_T)$. Therefore, the school has an incentive to implement an effort of $e_T (> e_V)$.

[7] For formal proofs, see Chakrabarti (2008b).

4. Data

The Florida data consist of school-level data on test scores, grades, socioeconomic characteristics of schools, and school finances; they are obtained from the Florida Department of Education. School-level data on test scores are obtained from the Florida Comprehensive Assessment Test. Mean scale scores (on a scale of 100-500) on grade 4 reading and grade 5 math are available for 1998-2002. Mean scale scores (on a scale of 1-6) on the Florida grade 4 writing test are available for 1994-2002.

Data on socioeconomic characteristics include sex composition (1994-2002), percentage of students eligible for free or reduced-price lunch (1997-2002), and race composition (1994-2002), and are obtained from the school indicators database of the Florida Department of Education. This study refers to school years by the calendar year of the spring semester. School finance data consist of several measures of school-level and district-level per-pupil expenditures, and are obtained from the school indicators database and the Office of Funding and Financial Reporting of the Florida Department of Education.

The Wisconsin data consist of school-level data on test scores, socioeconomic characteristics of schools, and per-pupil expenditures (both at the school and district levels). The data are obtained from the Wisconsin Department of Public Instruction, the Milwaukee Public Schools, and the Common Core of Data of the National Center for Education Statistics. School-level data on test scores are obtained for 1) the Third Grade Reading Test (renamed the Wisconsin Reading Comprehension Test, or WRCT, in 1996) and 2) the grade 5 Iowa Test of Basic Skills (ITBS). School scores for the WRCT, which was first administered in 1989, are reported in three "performance standard categories": percentage of students below, percentage of students at, and percentage of students above the standard.[8] Data for these three categories are available for 1989-97. School-level ITBS reading data (mean scores) are available for 1987-93; ITBS math data (mean scores) are available for 1987-97.

5. Empirical Strategy

5.1 Florida

In Florida, the schools that received an "F" grade in 1999 were directly exposed to the threat of vouchers because all their students would be eligible for vouchers if the school received another "F" in the next three years. These F-schools constitute the group of treated schools. Schools that received a "D" grade in 1999 were closest to the F-schools in terms of grade, but were not directly treated by the program. These "D-schools" constitute the group of control schools. The treatment and control groups consist of 65 and 457 elementary schools, respectively.[9] Because the program was announced in June 1999 and the grades were based on tests held in February 1999, we classify schools into treatment and control groups on the basis of their pre-program scores and grades.

If the F-schools and D-schools have similar trends in scores in the pre-program period, any shift of the F-schools compared with the D-schools in the post-program period can be attributed to the program.

The identifying assumption here is that if the F-schools and D-schools have similar trends in scores in the pre-program period, any shift of the F-schools compared with the D-schools in the post-program period can be attributed to the program. To test this assumption, we first run the following fixed-effects regression (and its ordinary least squares [OLS] counterpart) using only pre-program data:

$$(1) \qquad s_{it} = f_i + \alpha_0 t + \alpha_1 (F * t) + \alpha_2 X_{it} + \varepsilon_{it},$$

where s_{it} is the mean score of school i in year t, f_i are school-fixed effects, t denotes a time trend, F is a dummy variable taking a value of 1 for F-schools and 0 for D-schools, $F^* t$ is an interaction between the F dummy and trend, X_{it} denotes the set of school characteristics, and ε_{it} is a stochastic error term. Scores considered in the Florida part of the analysis include mean school scores in FCAT reading, FCAT math, and FCAT writing. The pre-program difference in trend of the F-schools is captured in α_1.

If F-schools and D-schools have similar pre-program trends, we investigate whether the F-schools demonstrate a higher improvement in test scores in the post-program era using specification 2 below. If the treated F-schools demonstrate a differential pre-program trend, then in addition to estimating this specification, we estimate a modified version in which we control for the pre-program differences in trends.

We estimate a completely unrestricted and nonlinear model that includes year dummies to control for common year effects and interactions of post-program year dummies with

[8] The method of reporting ITBS math and WRCT reading scores changed in 1998. Therefore, we use pre-1998 scores.

[9] We restrict our analysis to elementary schools because there were too few middle and high schools that received an "F" grade in 1999 (seven and five, respectively) to justify analysis.

the F-school dummy to capture individual post-program year effects:

$$(2) \quad s_{it} = f_i + \sum_{j=1999}^{2002} \beta_{0j} D_j + \sum_{j=1999}^{2002} \beta_{1j}(F * D_j) + \beta_2 X_{it} + \varepsilon_{it},$$

where D_j, $j = \{1999, 2000, 2001, 2002\}$ are year dummies for 1999, 2000, 2001, and 2002, respectively. While the above specification includes school-fixed effects, we also estimate an OLS counterpart to it. OLS regressions corresponding to both specifications 1 and 2 include a dummy for the treatment group *F*. Note that this is absorbed in the fixed-effects regressions because it is a time-invariant school effect.

Specification 2 does not constrain the post-program year-to-year gains of the F-schools to be equal and allows the program effect to vary across years. The coefficients β_{1t}, $i = 2000, 2001, 2002$ represent the effect of one, two, and three years into the program, respectively, for the F-schools. Given the nature of the Florida program, the 1999 threatened schools (that is, the schools that received an "F" grade in 1999) would be exposed to the threat of vouchers for the next three years only. Therefore, we track the performance of the threatened schools (relative to the control schools) for three years after the program—2000, 2001, and 2002—when the threat of vouchers would be in effect.

The above specifications assume that the D-schools were not affected by the program. Although the D-schools did not face any direct threat from the program, they might have faced an indirect threat because they were close to receiving an "F" grade.[10] Therefore, we next allow the F-schools and D-schools to be different treated groups (with varying intensities of treatment) and compare their post-program improvements, if any, with 1999 "C-schools," which are the next grade up in the scale using the above specifications after adjusting for another treatment group. It should be noted that since D-schools and C-schools may face the threat to some extent, our estimates may be underestimates (lower bounds), but not overestimates.

5.2 Milwaukee

Our strategy is based on and is similar to that of Hoxby (2003b). Since students in the Milwaukee Public Schools eligible for free or reduced-price lunch were also eligible for vouchers, the extent of treatment of the Milwaukee schools depended on the percentage of students eligible for free or reduced-price lunch.[11] Using this information, Hoxby classifies the Milwaukee schools into two treatment groups based on the percentages of students eligible for free or reduced-price lunch—"most treated" (at least two-thirds of students eligible in the pre-program period) and "somewhat treated" (less than two-thirds of students eligible in the pre-program period).

We classify the schools into three treatment groups (in contrast to Hoxby's two) based on their pre-program (1989-90 school year) percentage of students eligible for free or reduced-price lunch. Thus, our treatment groups are more homogenous as well as starker from each other. Additionally, to test the

Since students in the Milwaukee Public Schools eligible for free or reduced-price lunch were also eligible for vouchers, the extent of treatment of the Milwaukee schools depended on the percentage of students eligible for free or reduced-price lunch.

robustness of our results, we consider alternative samples obtained by varying the cutoffs that separate the different treatment groups, departing from the Hoxby approach. The 60-47 (66-47) sample classifies schools that have at least 60 percent (66 percent) of students eligible for free or reduced-price lunch as "more treated," schools with such population between 60 percent (66 percent) and 47 percent as "somewhat treated," and schools with such population less than 47 percent as "less treated." We also consider alternative classifications, such as "66" and "60" samples, where there are two treatment groups—schools that have at least 66 percent (60 percent) of students eligible for free or reduced-price lunch are designated as more treated schools, and schools with such population below 66 percent (60 percent) as somewhat treated schools. Since there were very few middle and high schools in the Milwaukee Public Schools and student participation in the Milwaukee Parental Choice Program was mostly in the elementary grades, we restrict our analysis to elementary schools.

[10] In fact, there is some anecdotal evidence that D-schools may have responded to the program. The superintendent of Hillsborough County, which had no F-schools in 1999, announced that he would take a 5 percent pay cut if any of his thirty-seven D-schools received an "F" grade on the next school report card. For more information, see Innerst (2000).

[11] Under the Milwaukee program, all households at or below 175 percent of the poverty line are eligible to apply for vouchers. Households at or below 185 percent of the poverty line are eligible for free or reduced-price lunch. However, the cutoff of 175 percent is not strictly enforced (Hoxby 2003a), and households within this 10 percent margin are often permitted to apply. In addition, there were very few students who fell in the 175 percent-185 percent range, while in fact 90 percent of students eligible for free or reduced-price lunch qualified for free lunch (Witte 2000). Students below 135 percent of the poverty line qualified for free lunch.

Our control group criteria are also based on Hoxby (2003b). Since all schools in Milwaukee were potentially affected by the program, Hoxby constructs a control group that consists of Wisconsin schools outside Milwaukee that satisfy the following criteria in the pre-program period that: 1) had at least 25 percent of their population eligible for free or reduced-price lunch, 2) had black students who make up at least 15 percent of the population, and 3) were urban. Her control group consists of twelve schools.

For our control schools, we designate schools that are located outside Milwaukee but within Wisconsin, satisfy Hoxby's first two criteria, and have locales as similar as possible to the Milwaukee schools. Note that all of these characteristics pertain to the pre-program school year 1989-90.[12]

Using each sample, we investigate how the different treatment groups in Milwaukee responded to the "voucher shock" program. Using specification 3 below, we first test whether the pre-program trends of the untreated and the different treated groups were the same. We then estimate OLS and fixed-effects versions of specification 4 below. If we observe differences in pre-existing trends between the different treated groups of schools, then in addition to estimating specification 4, we estimate modified versions of the specification that control for pre-existing differences in trends:

$$(3) \qquad s_{it} = f_i + \gamma_0 t + \sum_k \gamma_{1k} (I_k * t) + \gamma_2 X_{it} + \varepsilon_{it}$$

$$(4) \qquad s_{it} = f_i + \sum_{j=1989}^{2007} \delta_{0j} D_j + \sum_{j=1989}^{2007} \delta_{1kj} (I_k * D_j) + \gamma_2 X_{it} + \varepsilon_{it} ,$$

where s_{it} denotes scores of school i in period t; D_j, $j = \{1989, \ldots, 2007\}$ are year dummies for 1989 through 2007, respectively; $k \in \{MT, ST, LT\}$ for the WRCT and $k \in \{MT, ST\}$ for the ITBS, where MT denotes "more treated," ST denotes "somewhat treated," and LT denotes "less treated." The scores considered are mean scores in ITBS reading and ITBS math as well as percentages of students above the standard in WRCT reading.

6. Results

Table 2 presents baseline characteristics of treated and control groups in Florida and Wisconsin. It shows that the more treated schools in Florida were indeed similar to the more treated schools in Wisconsin and, except in one case, the differences between them were not statistically significant. Similarly, the control schools in Florida were similar to the control schools in Wisconsin, and the differences between them were not statistically significant.

However, the treated schools were somewhat different from the control schools within each state. The reason is that Wisconsin schools outside Milwaukee were considerably more advantaged than schools in Milwaukee. We arrived at this control group despite using the strategy (following Hoxby [2003a, b]) of selecting control schools as similar as possible to Milwaukee's more treated schools in terms of pre-program characteristics.

It is important that both the more treated schools and the control groups be similar across the two programs in terms of pre-program characteristics as well as across the two locations. As a result, for purposes of comparing effects across the two

[12] The more treated and control group characteristics are presented in Table 2. In the 66-47 sample, the somewhat treated (less treated) group had an average of 55.4 percent (37.17 percent) of students eligible for free or reduced-price lunch, 50.99 percent (45.37 percent) who were black, and 4.09 percent (3.83 percent) who were Hispanic.

Table 2

Pre-Program Demographic Characteristics of Florida and Wisconsin More Treated and Control Schools
Percent

Panel A: More Treated Schools

		Wisconsin		Florida–Wisconsin	
	Florida	66-47	60-47	66-47	60-47
Black	62.79	66.55	62.90	-3.76	-0.10
	(28.23)	(32.22)	(29.58)	[0.56]	[0.99]
Hispanic	18.95	18.07	14.81	0.88	4.14
	(23.40)	(24.54)	(21.86)	[0.87]	[0.36]
White	17.18	10.21	17.38	6.97	-0.20
	(19.54)	(10.68)	(16.55)	[0.07]	[0.96]
Male	51.38	52.25	52.33	-0.87	-0.95
	(4.84)	(2.60)	(2.58)	[0.34]	[0.22]
Free or reduced-price lunch	85.80	84.5	82.9	1.3	2.9
	(9.95)	(6.48)	(9.04)	[0.50]	[0.12]

Panel B: Control Schools

	Florida	Wisconsin	Florida–Wisconsin
Black	18.12	22.37	-4.25
	(14.17)	(12.93)	[0.10]
Hispanic	15.49	14.84	0.17
	(21.23)	(6.02)	[0.86]
White	63.59	60.85	2.73
	(22.33)	(12.80)	[0.49]
Male	51.38	50.63	0.76
	(4.84)	(2.29)	[0.43]
Free or reduced-price lunch	50.14	44.95	5.19
	(17.51)	(11.66)	[0.10]

Source: Author's calculations.

Notes: The group of Florida more treated and control schools is composed of F-schools and C-schools, respectively. Samples 66-47 and 60-47 are described in Section 5.2 of the article. Standard deviations are in parentheses; p-values are in brackets.

programs, we use the C-schools in Florida as the control group. Noticeably, the control group in Wisconsin was very similar to the C-schools in Florida and was not statistically different from them in terms of any characteristics (Table 2). Still another reason for selecting the C-schools as the control group in Florida was that while the D-schools were more similar to the more treated F-schools in terms of grade and demographics, they were very close to receiving an "F" grade; hence, to some extent they perceived an indirect threat and to some extent were treated by the program.

Because of differences between the treated and control schools, one might argue that in the absence of the program, the control group would have evolved differently from the more treated group. However, multiple years of pre-program data allow us to check (and control) for any differences in pre-program trends of these groups. In this way, we can dispose of any level differences between the treated and control groups as well as control for differences in pre-program trends, if any. It seems likely that once we control for differences in trends as well as in levels, any remaining differences between the treated and control groups will be minimal. In other words, our identifying assumption is that if the treated schools followed the same trends as the control schools in the immediate pre-program period, they would have evolved similarly in the immediate post-program period in the absence of the program. We also control for time-varying observable characteristics. School-fixed effects remove any time-invariant unobservable characteristics. Note that while time-varying unobserved characteristics cannot be directly controlled for, they did not drive the results as long as the F-schools did not experience a differential shock in unobserved characteristics that coincided with the timing of the program.

6.1 Florida

Considerable anecdotal evidence suggests that F-schools have responded to the voucher program. Just after the program's inception, Escambia County implemented a 210-day extended school year in its F-schools (the typical duration was 180 days), introduced an extended school day at least twice a week, and added small-group tutoring on afternoons and Saturdays and longer time blocks for writing and math instruction. To curb absenteeism, the county started an automated phone system to contact parents when a child is absent. Miami-Dade County hired 210 additional teachers for its twenty-six F-schools, switched to phonics instruction, and encouraged parents (many of whom were dropouts) to go back to school for a high-school-equivalency diploma. Broward County reduced its class size to eighteen to twenty students in its low-performing schools and increased services for children whose primary language is not English. Palm Beach County targeted its fourth-grade teachers for coaching and began more frequent and closer observation of teachers in its F-schools (Innerst 2000). Carmen Varela-Russo, Associate Superintendent of Technology, Strategic Planning, and Accountability at Broward County Public Schools, described the situation this way: "People get lulled into complacency . . . the possibility of losing children to private schools or other districts was a strong message to the whole community" (Innerst 2000). The analysis below investigates whether the data in Florida support this behavior.

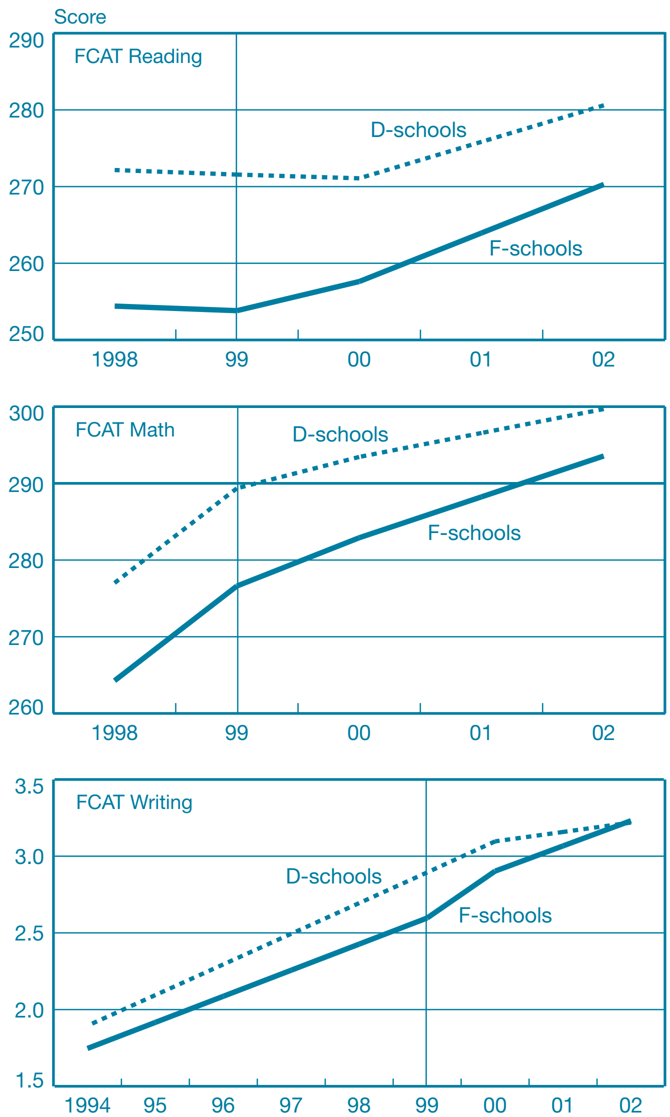
> *Considerable anecdotal evidence suggests that F-schools have responded to [Florida's] voucher program.*

Chart 2, which depicts trends in reading, math, and writing scores in F-schools and D-schools, shows that 1999 was the watershed year. In both reading and math, the F-schools had similar trends before the program. However, the F-schools showed improvement relative to the D-schools after the program, and the gap between F- and D-schools narrowed. In writing, while the F-schools were deteriorating relative to the D-schools before the program, this pattern changed after it. The F-schools showed improvement relative to the D-schools to the extent that they successfully closed the "F" to "D" gap after the program.

We now turn to our estimation results. All regressions control for ethnicity (the percentage of students in different racial categories in a school), the percentage of male students, the percentage of students eligible for free or reduced-price lunch, and real per-pupil expenditures. Table 3 presents pre-program trends in reading, math, and writing. It reveals that F-schools have no significant differences in trend compared with D-schools in reading and math, although they exhibit a small, negative differential trend in writing. Compared with C-schools, F-schools exhibit a negative differential trend in reading and writing, but no significant differential trend in math. D-schools exhibit a negative trend in reading and a positive trend in math and writing compared with C-schools. Whenever there is a difference in pre-program trends, our reported regressions control for these differences by including interactions between trend and the respective treated dummies.

Chart 2

Florida "Threat-of-Voucher" Program



Source: Author's calculations.

Note: FCAT is the Florida Comprehensive Assessment Test.

Table 4, columns 1-3 present the effects of the Florida threat-of-voucher program on F-school reading, math, and writing scores compared with those for D-schools. All models reported include school-fixed effects. The results from our OLS estimation are similar to the fixed-effects estimates and hence are not reported. The regressions for writing include interactions of the "F" dummy with trend to control for differences in pre-program trends seen above.[13,14] The table shows economically large, positive, and statistically significant effects in each subject area and year.

D-schools are considered as an additional treatment group in Table 4, columns 4-6. Here, we see how the program affects F-schools (more treated) and D-schools (less treated) compared with C-schools. All columns control for differences in pre-existing trends between groups. The results show positive, significant year effects in reading, math, and writing for F-schools in each of the years after the program's implementation. Although many of the D-school effects are also positive and significant, the F-school shifts are statistically larger in each year.[15] The F-school effects are economically meaningful as well. In reading, relative to the base year, F-schools showed a 3.6 percent improvement in the first year after the program, a 4.2 percent improvement after the second year, and a 6.3 percent improvement after the third year. In math, F-schools showed a 3.4 percent, 4.2 percent, and 4.5 percent improvement in the first, second, and third years, respectively, after implementation of the program. In writing, the percentage improvement was around 15 percent. At the end of 2002 (three years after program implementation), the pre-program gap between F-schools and C-schools was closed by 37.08 percent in reading, 30.31 percent in math, and around 75 percent in writing.

Our results show considerable improvement in the F-schools after the program compared with the control schools.

In summary, based on different samples, different subjects, and both OLS and fixed-effects estimates, our results show considerable improvement in the F-schools after the program compared with the control schools. Although D-schools show non-negligible improvement (at least in reading and writing), their improvement is considerably less than and statistically different from that of the F-schools.

[13] Note that the table reports only the coefficients that reflect program effects; therefore, the coefficient corresponding to this interaction term (which captures the differential pre-existing trend) is not reported. Pre-existing trends are reported in Table 3.

[14] The regressions for reading and math (columns 1 and 2) do not include this interaction term because there is no evidence of differential pre-program trends in reading and math for F-schools and D-schools (Table 3). Note that the results with inclusion of this term remain very similar.

[15] Here, we test whether the F-school effects are statistically different from the D-school effects against the null hypothesis that they are equal.

Table 3

Pre-Program Trend of F-, D-, and C-Schools in Florida

	Sample of F- and D-Schools						Sample of F-, D-, and C-Schools					
	FCAT Reading		FCAT Math		FCAT Writing		FCAT Reading		FCAT Math		FCAT Writing	
	OLS (1)	FE (2)	OLS (3)	FE (4)	OLS (5)	FE (6)	OLS (7)	FE (8)	OLS (9)	FE (10)	OLS (11)	FE (12)
Trend	0.41	-0.05	13.20***	13.02**	0.20**	0.21**	2.66**	2.70	9.79***	10.20***	0.18***	0.19***
	(0.56)	(0.47)	(0.55)	(0.61)	(0.008)	(0.003)	(0.57)	(0.36)	(0.53)	(0.38)	(0.01)	(0.002)
F * trend	-1.78	-2.01	-0.98	-0.72	-0.05***	-0.04***	-3.80	-4.77***	2.46	1.96	-0.03***	-0.03***
	(2.47)	(1.46)	(1.44)	(1.48)	(0.011)	(0.007)	(2.29)	(1.41)	(1.51)	(1.43)	(0.01)	(0.01)
D * trend							-2.29***	-2.69***	3.46***	2.79***	0.02**	0.02***
							(0.66)	(0.57)	(0.60)	(0.67)	(0.007)	(0.003)
Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	1,013	1,013	1,006	1,006	2,948	2,948	2,386	2,386	2,377	2,377	6,982	6,982
R^2	0.58	0.93	0.59	0.90	0.64	0.80	0.76	0.95	0.74	0.93	0.65	0.82

Source: Author's calculations.

Notes: FCAT is the Florida Comprehensive Assessment Test; OLS is ordinary least squares regression; FE is fixed-effects regression. Controls include race, sex, percentage of students eligible for free or reduced-price lunch, and real per-pupil expenditure. Huber-White standard errors are in parentheses. All regressions are weighted by the number of students tested.

***Statistically significant at the 1 percent level.
**Statistically significant at the 5 percent level.
*Statistically significant at the 10 percent level.

Table 4

Effect of "Threatened Status" on FCAT Reading, Math, and Writing Scores

Sample of Treated F- and Control D-Schools in Florida

	Reading FE (1)	Math FE (2)	Writing FE (3)	Reading FE (4)	Math FE (5)	Writing FE (6)
Treated * one year after program	4.85***	6.78***	0.35***			
	(1.68)	(1.63)	(0.04)			
Treated * two years after program	3.30*	7.25***	0.37***			
	(1.71)	(1.82)	(0.04)			
Treated * three years after program	7.08***	5.35***	0.43			
	(1.78)	(2.00)	(0.05)			
Less treated * one year after program				3.53***	0.97	0.05**
				(0.76)	(0.85)	(0.02)
Less treated * two years after program				5.52***	2.54***	0.00
				(0.80)	(0.94)	(0.02)
Less treated * three years after program				7.94***	3.47***	-0.03
				(0.87)	(0.92)	(0.02)
More treated * one year after program				9.32[b]***	8.96[b]***	0.39[b]***
				(1.87)	(1.59)	(0.04)
More treated * two years after program				10.75[a]***	11.00[b]***	0.37[a]***
				(1.87)	(1.77)	(0.04)
More treated * three years after program				16.03[b]***	11.94[b]***	0.39[a]***
				(1.91)	(1.95)	(0.05)
School-fixed effects	Y	Y	Y	Y	Y	Y
Year dummies	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y
Observations	2,550	2,524	4,476	5,933	5,909	10,587
R^2	0.77	0.76	0.85	0.86	0.83	0.86
p-value[c]	0.00	0.00	0.00	0.00	0.00	0.00

Source: Author's calculations.

Notes: FCAT is the Florida Comprehensive Assessment Test. FCAT scores for reading and math are for the period 1998-2000; FCAT scores for writing are for the period 1994-2002. FE is fixed-effects regression. Huber-White standard errors are in parentheses. Controls include race, sex, percentage of students eligible for free or reduced-price lunch, and real per-pupil expenditure. All regressions are weighted by the number of students tested.

[a] More treated significantly different from less treated at 5 percent level.
[b] More treated significantly different from less treated at 1 percent level.
[c] *p*-value of *F*-test of the program effect on treated schools.

***Statistically significant at the 1 percent level.
**Statistically significant at the 5 percent level.
*Statistically significant at the 10 percent level.

Table 5
Pre-Program Trend of More Treated, Somewhat Treated, and Less Treated Schools in Milwaukee

	WRCT (Percentage above)		ITBS Reading		ITBS Math	
	OLS	FE	OLS	FE	OLS	FE
	(1)	(2)	(3)	(4)	(5)	(6)
Trend	-3.84	-4.34**	-4.09	-3.45	-3.04*	2.52**
	(2.33)	(2.16)	(4.11)	(3.42)	(1.66)	(0.98)
More treated * trend	-3.08	-2.03	4.01	-1.88	0.56	0.32
	(3.41)	(3.35)	(3.69)	(2.73)	(1.97)	(1.40)
Somewhat treated * trend	-4.41	-3.61	3.14	2.12	0.73	0.31
	(3.01)	(2.67)	(4.05)	(3.17)	(1.83)	(1.21)
Less treated * trend	-2.33	-3.23				
	(3.61)	(3.10)				
Observations	242	242	411	411	410	410
R^2	0.50	0.87	0.30	0.56	0.30	0.71

Source: Author's calculations.

Notes: WRCT is the Wisconsin Reading Comprehension Test; ITBS is the Iowa Test of Basic Skills; OLS is ordinary least squares regression; FE is fixed-effects regression. Controls include race, sex, and percentage of students eligible for free or reduced-price lunch.

***Statistically significant at the 1 percent level.
**Statistically significant at the 5 percent level.
*Statistically significant at the 10 percent level.

6.2 Milwaukee

The Milwaukee analysis uses the 66-47 sample. Estimation results for pre-program trends are presented in Table 5. The results show no statistical difference in trends between the various treated and control groups in any subject area.

Table 6 examines the effect of the Milwaukee "voucher shock" program on the WRCT (the percentage above), ITBS reading, and ITBS math scores of different treated groups. Except for the positive and statistically significant effect in WRCT reading in the test's second year, there is no statistically significant evidence of any effect of the program. Although the second year's somewhat treated effect in ITBS math is statistically significant, it is more than the corresponding more treated effect.[16]

Thus, the results in Milwaukee are mixed. The program seems to have had a positive and significant effect in the second year after the program's implementation, at least in the WRCT.

[16] Since the ITBS was administered in Milwaukee as a district assessment program, we do not have data on non–Milwaukee, Wisconsin, schools for this test. As a result, our comparison group is the less treated group of schools. Since the comparison group is also treated to some extent, we expect our estimates for the ITBS to be lower bounds.

These results seem to be robust in that they are replicated in the analysis with other samples.[17] Chart 3 presents the trends in ITBS scores for the various groups. As expected, there is no evidence of any program effect.

The results show no statistical difference in trends between the various treated and control groups in any subject area. . . . Except for the positive and statistically significant effect in [Wisconsin Reading Comprehension Test] reading in the test's second year, there is no statistically significant evidence of any effect of the program. . . . Thus, the results in Milwaukee are mixed.

[17] These results are not reported here, but are available from the author.

Table 6

Effect of the Milwaukee "Voucher Shock" Program

	WRCT (1)	ITBS Reading (2)	ITBS Math (3)
Somewhat treated * one year after program	2.03	4.15	-1.35
	(2.81)	(4.49)	(2.94)
Somewhat treated * two years after program	5.38**	7.83	6.14*
	(2.43)	(5.17)	(3.38)
Somewhat treated * three years after program	5.01	6.78	2.47
	(3.03)	(5.31)	(3.31)
More treated * one year after program	-0.92	1.12	-4.02
	(3.33)	(3.86)	(3.26)
More treated * two years after program	6.06*	6.59	4.36
	(3.14)	(5.15)	(3.83)
More treated * three years after program	5.69	2.85	-2.22
	(3.98)	(5.18)	(3.54)
School-fixed effects	Y	Y	Y
Year dummies	Y	Y	Y
Controls	Y	Y	Y
Observations	1,195	717	1,127
R^2	0.58	0.55	0.60
p-value[a]	0.11	0.62	0.27

Source: Author's calculations.

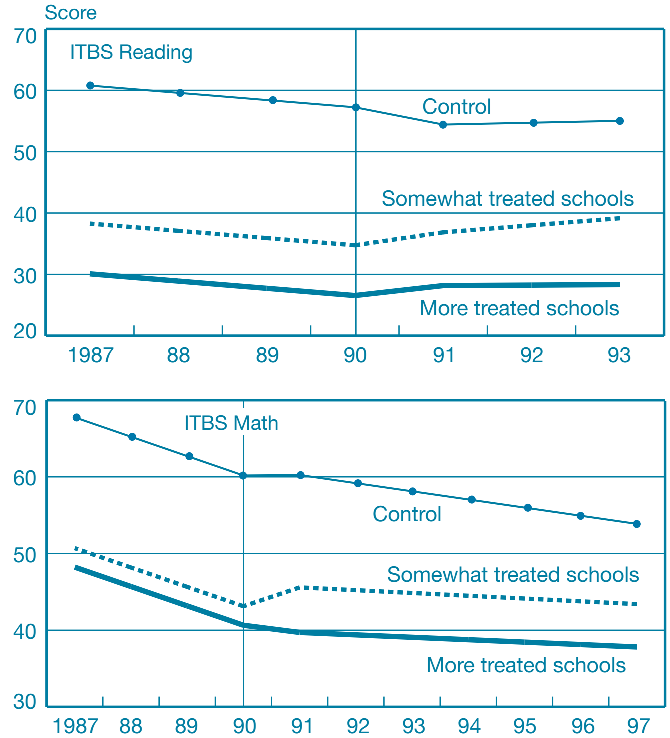
Notes: WRCT is the Wisconsin Reading Comprehension Test; ITBS is the Iowa Test of Basic Skills. Huber-White standard errors are in parentheses. All regressions include school-fixed effects and control for race, sex, percentage of students eligible for free or reduced-price lunch, and real per-pupil expenditure.

[a] *p*-value of the F-test of joint significance of more treated shift coefficients.

***Statistically significant at the 1 percent level.
**Statistically significant at the 5 percent level.
*Statistically significant at the 10 percent level.

Chart 3

Milwaukee "Voucher-Shock" Program



Source: Author's calculations.

Note: ITBS is the Iowa Test of Basic Skills.

7. Robustness Checks

7.1 Mean Reversion

Several factors might bias the results; we consider each factor and its potential solutions. First is the issue of mean reversion. Mean reversion is the statistical tendency whereby high- or low-scoring schools tend to score closer to the mean subsequently. Because the F-schools scored low in 1999, a natural question would be whether the improvement in Florida is driven by mean reversion rather than the voucher program. Since we conduct a difference-in-differences analysis, our estimates will be tainted by mean reversion only if F-schools mean-revert to a greater extent than do the D-schools or the C-schools, or both.

To investigate mean reversion, we examine whether and by how much schools that received an "F" grade in 1998 improved during the 1998-99 academic year compared with those that received a "D" (or "C") grade in 1998. Since these years fall within the pre-program period, the gain can be taken to approximate the mean-reversion effect and can be subtracted from the post-program gain of F-schools compared with D-schools (or C-schools) to get at the mean-reversion-corrected program effect.

The accountability system of assigning letter grades to schools began in 1999. The pre-1999 accountability system classified schools into four groups, designated 1 (low) to 4 (high). However, using the state grading criteria and data on the percentage of students in different achievement levels in each FCAT reading, math, and writing, we assigned letter grades to schools in 1998 and implemented the above strategy. Schools receiving "F," "D," and "C" grades in 1998 using this procedure are referred to as "98F-schools," "98D-schools," and "98C-schools," respectively.

Using Florida data for 1998 and 1999, we demonstrate in Table 7, panel A, that when compared with the 98D-schools, the 98F-schools show no evidence of mean reversion either in reading or math, although there is mean reversion in writing. Compared with the 98C-schools (panel B), there is no evidence of mean reversion in reading; both 98D-schools and 98F-schools show comparable amounts of mean reversion in math; only 98F-schools show mean reversion in writing.

Table 7

Mean Reversion of 98F-Schools Compared with 98D- and 98C-Schools, 1998-99

Panel A: 98F- and 98D-Schools

	Dependent Variable: FCAT Score, 1998-99		
	Reading FE (1)	Math FE (2)	Writing FE (3)
Trend	2.01***	14.02***	0.04***
	(0.43)	(0.49)	(0.01)
98F * trend	-0.65	1.17	0.14***
	(1.14)	(1.19)	(0.02)
Observations	1,353	1,354	1,355
R^2	0.93	0.91	0.85

Panel B: 98F-, 98D-, and 98C-Schools

	Dependent Variable: FCAT Score, 1998-99		
	Reading FE (1)	Math FE (2)	Writing FE (3)
Trend	1.76***	9.71***	0.03***
	(0.35)	(0.36)	(0.01)
98F * trend	-0.55	4.63***	0.14***
	(1.12)	(1.16)	(0.02)
98D * trend	0.16	4.22***	0.01
	(0.54)	(0.58)	(0.01)
Observations	2,605	2,608	2,608
R^2	0.96	0.94	0.87

Source: Author's calculations.

Notes: FCAT is the Florida Comprehensive Assessment Test; FE is fixed-effects regression. All regressions control for race, sex, percentage of students eligible for free or reduced-price lunch, and real per-pupil expenditure. The ordinary least squares regressions include 98F- and 98D-school dummies. In the sample of 98F- and 98D-schools, the standard deviations of FCAT reading, math, and writing are 18.9, 18.05, and 0.30, respectively. In the sample of 98F-, 98D-, and 98C-schools, the standard deviations of FCAT reading, math, and writing are 21.16, 21.56, and 0.31, respectively.

***Statistically significant at the 1 percent level.
**Statistically significant at the 5 percent level.
*Statistically significant at the 10 percent level.

Table 8

Is There a Stigma Effect of Getting the Lowest Performing Grade?
Effect of Being Categorized in Group 1 on FCAT Writing Scores

	Using FCAT Writing Scores, 1997-98					
	Sample: Group 1, 2 Schools			Sample: Group 1, 2, 3 Schools		
	OLS (1)	FE (2)	FE (3)	OLS (4)	FE (5)	FE (6)
Trend	0.52***	0.52***	0.48***	0.48***	0.48***	0.46***
	(0.04)	(0.03)	(0.04)	(0.02)	(0.01)	(0.02)
Group 1 * trend	-0.01	-0.02	-0.02	0.03	0.01	0.02
	(0.08)	(0.06)	(0.06)	(0.07)	(0.05)	(0.05)
Group 2 * trend				0.03	0.04	0.04
				(0.04)	(0.03)	(0.03)
Controls	N	N	Y	N	N	Y
Observations	314	314	314	1,361	1,361	1,358
R^2	0.49	0.84	0.85	0.52	0.87	0.87

Source: Author's calculations.

Notes: FCAT is the Florida Comprehensive Assessment Test; OLS is ordinary least squares regression; FE is fixed-effects regression. Huber-White standard errors are in parentheses. All regressions are weighted by the number of students tested; controls include race, sex, percentage of students eligible for free or reduced-price lunch, and real per-pupil expenditure. The OLS regressions include group 1 and group 2 dummies.

***Statistically significant at the 1 percent level.
**Statistically significant at the 5 percent level.
*Statistically significant at the 10 percent level.

7.2 Stigma Effect of Getting the Lowest Performing Grade

A second concern in Florida is the potential stigma effect of receiving a performance grade of "F." If there is such a stigma, the F-schools will try to improve only to avoid this stigma rather than in response to the program. We use several alternative strategies to investigate this possibility. First, although the system of assigning letter grades to schools started in 1999, Florida had an accountability system in the pre-1999 period when schools were categorized into four groups, designated 1 (low) to 4 (high), based on FCAT writing and reading and math norm-referenced test scores. Using FCAT writing data for two years (1997 and 1998), we investigate whether the schools, which were categorized in group 1 in 1997, improved in relation to the 1997 group 2 and group 3 schools in 1997-98.[18] Our rationale is that if a stigma effect is associated with getting the lowest performing grade, the group 1 schools should improve relative to the group 2 and 3 schools, even in the absence of the threat-of-voucher program.

If there is [a low-performance] stigma, the F-schools will try to improve only to avoid this stigma, rather than in response to the program.

Table 8, using pre-program FCAT writing scores, shows that no such stigma effect exists—group 1 schools display no improvement relative to the group 2 or group 3 schools.

Second, all the schools that received an "F" grade in 1999 received higher grades in 2000, 2001, and 2002. Therefore, although the stigma effect on F-schools may be operative in 2000, this is not likely to be the case in 2001 or 2002 since none of the F-schools received an "F" grade in the preceding year (2000 or 2001, respectively). However, the F-schools would face the threat of vouchers until 2002, so any improvement in

[18] We do not use the pre-1999 reading and math norm-referenced test (NRT) scores because different districts used different NRTs during this period, which varied in content and norms. Also, districts often chose different NRTs in different years. Thus, these NRTs were not comparable across districts and across time. Moreover, since districts could choose the specific NRT to administer each year, the choice was likely related to time-varying (and also time-invariant) district-unobservable characteristics that also affected test scores.

2001 and 2002 would provide evidence in favor of the threat-of-voucher effect and against the stigma effect. F-schools showed strong gains in both 2001 and 2002—a result that provides further support for the threat-of-voucher effect and against the stigma effect.

7.3 Sorting

Another factor relates to sorting in the context of the Milwaukee voucher program. Vouchers affect public school quality not only through direct public school response but also through changes in student composition and peer quality brought about by sorting. These three factors are then reflected in the public school scores.[19] This issue is important in Milwaukee because over the years students have left the city's public schools with vouchers. In contrast, no Florida school became eligible for vouchers in 2000 or 2001. Therefore, the program effects (for each of the years 2000, 2001, and 2002) are not likely to be tainted by this factor.[20] Moreover, as we discuss shortly, the demographic compositions of the different groups of schools remained very similar across the years under consideration.

We also examine whether the demographic composition of the different Milwaukee treated groups changed over the years (Table 9). No such evidence is found. Only a few of the coefficients are statistically significant, and they are always very small in magnitude. They imply changes of less than 1 percent, more precisely, ranging between 0.22 percent and 0.65 percent. This result suggests that sorting was not an important factor. Note that we conducted the same exercise for Florida as well and found no evidence of any relative shift of the demographic composition of the F-schools compared with the D-schools or C-schools.

Vouchers affect public school quality not only through direct public school response but also through changes in student composition and peer quality brought about by sorting.

[19] See Hsieh and Urquiola (2006) for a discussion.

[20] This does not mean that the Florida program was not credible. Ten schools received a second "F" grade in 2002, nine schools in 2003, and twenty-one in 2004; all of these students became eligible for vouchers.

Table 9
Effect of Milwaukee Program on Demographic Composition of Schools
Percent

	Black (1)	Hispanic (2)	Asian (3)
Less treated * program	0.90	0.40	0.04
	(1.59)	(0.83)	(0.37)
Somewhat treated * program	-0.25	1.06	0.53
	(1.35)	(0.63)	(0.37)
More treated * program	-1.0	1.57	0.65*
	(1.34)	(0.81)	(0.37)
Less treated * program * trend	0.22	0.16	0.24***
	(0.32)	(0.15)	(0.07)
Somewhat treated * program * trend	0.70	-0.12	0.29***
	(0.25)	(0.13)	(0.07)
More treated * program * trend	0.08	-0.39***	-0.22***
	(0.23)	(0.14)	(0.07)
Observations	1,228	1,226	1,216
R^2	0.95	0.97	0.91

Source: Author's calculations.

Notes: Huber-White standard errors are in parentheses. All regressions are weighted by the number of students tested. All columns include a time trend, a program dummy that takes a value of 1 after the program, and an interaction between program dummy and trend.

***Statistically significant at the 1 percent level.
**Statistically significant at the 5 percent level.
*Statistically significant at the 10 percent level.

A Comparison of Program Effects in Florida and Milwaukee

Since Florida and Milwaukee are in different regions, we argue that our comparison of the effects of the two programs is fair and reasonable. First, as mentioned earlier, apart from the crucial design differences between the two programs, the other features of the programs were very similar. In both programs, private schools could not discriminate against voucher applicants. Also, the method of funding for the two programs, the average voucher amounts, and the per-pupil revenue losses from vouchers were very similar. Second, state and local revenues constituted very similar proportions of total revenue during the relevant periods—the percentages of revenue from state and local sources were 51 percent and 41 percent, respectively, in Florida, and 55 percent and 36 percent, respectively, in Milwaukee. Third, the demographic characteristics of the more treated and control schools in Florida were very similar, both economically and statistically, to those of the more treated and control schools in Milwaukee

Table 10

Comparison of Results from Florida "Threat-of-Voucher" and Milwaukee "Voucher-Shock" Programs Using Standardized Reading and Math Scores

					Corrected for Mean Reversion			
	Reading		Math		Reading		Math	
	Wisconsin WRCT (1)	Florida FCAT (2)	Wisconsin ITBS (3)	Florida FCAT (4)	Wisconsin WRCT (5)	Florida FCAT (6)	Wisconsin ITBS (7)	Florida FCAT (8)
More treated * one year after program	-0.06	0.47***	-0.24	0.45***	-0.06	0.47***	-0.24	0.24***
More treated * two years after program	0.38*	0.50***	0.26	0.55***	0.38*	0.50***	0.26	0.34***
More treated * three years after program	0.35	0.80***	-0.13	0.60***	0.36	0.80***	-0.13	0.39***

Source: Author's calculations.

Notes: Reading test scores are from the Wisconsin Reading Comprehensive Test (WRCT), 1989-97, and the Florida Comprehensive Assessment Test (FCAT) Reading, 1998-2002. Math test scores are from the Iowa Test of Basic Skills (ITBS) Math, 1986-97, and the FCAT Math, 1998-2002. All figures are respective sample standard deviations. All figures are obtained from regressions that contain school-fixed effects, year dummies, interactions of year dummies with the respective treatment dummies, race, sex, percentage of students eligible for free or reduced-price lunch, and real per-pupil expenditure. Standard deviation of FCAT reading scores = 20; standard deviation of FCAT math scores = 20; standard deviation of WRCT (percentage above) reading scores = 16; standard deviation of ITBS reading scores = 18.45; standard deviation of ITBS math scores = 16.71. For standard deviations corresponding to the mean reversion sample, see the notes to Table 4.

***Statistically significant at the 1 percent level.

**Statistically significant at the 5 percent level.

*Statistically significant at the 10 percent level.

(Table 2). Fourth, we repeat our analysis by comparing the improvement in Milwaukee with that of a large urban district in Florida: Miami-Dade County (the state's largest school district). The results are very similar and hence are not reported here. Finally, and perhaps most importantly, since we follow a difference-in-differences strategy in trends, any level or even trend differences between the two regions (that are common to schools in that region) are differenced out. It is unlikely that any remaining difference, which differentially affects the *trends* in the two regions *only* in the post-program period, will be large.

Table 10 compares the effects of the Florida and Milwaukee programs on their respective more treated schools both before and after correcting for mean reversion. Figures are based on data in Tables 4 and 6, and all numbers are expressed in terms of their respective sample standard deviations. Columns 1-4 present results before correcting for mean reversion; columns 5-8 present results corrected for mean reversion. Pre-correction results show positive and significant effect sizes in each of the years and subject areas in Florida, which always exceed the corresponding Milwaukee effect sizes (which are not significant, except in second-year reading). Mean-reversion-corrected effect sizes are obtained by subtracting the effect size attributed to mean reversion (obtained from expressing the relevant coefficients in Table 7, panel B, in terms of respective standard deviations) from the F-school effect sizes (obtained from expressing the more treated coefficients in Table 4, columns 4-6, in terms of respective sample standard deviations) in each of the three years after the program. The estimates in reading are the same as those described earlier. In math, although the effect sizes fall in Florida, they are still positive and considerably larger than those in Milwaukee. In reading (math), relative to the control schools, the F-schools show an improvement of 0.47 (0.24) standard deviations in the first year after the program, 0.5 (0.34) standard deviations after the second year, and 0.8 (0.39) standard deviations after the third year. Mean-reversion-corrected effect sizes in writing are 0.29, 0.25, and 0.29 in the first, second, and third years, respectively, after the program. Note that since none of the F-schools received an "F" grade in either 2000 or 2001, the mean-reversion-corrected effect sizes attributed to the Florida program in the second and third years may be underestimates.

8. Lessons for New York City

Our analysis of school voucher programs implies that policies that threaten underperforming public schools (or other agents) with functional and credible sanctions can induce them to respond in a way intended or desired by the policymaker. This finding has important implications for some educational policies in New York City. These include New York City's own accountability policy, also known as the "Progress Report" policy, and the federal No Child Left Behind law, as implemented by New York State.

The Progress Report policy was introduced in New York City in 2007. It rates schools on a scale of A to F, with grades based on three components: school environment, student performance, and student progress. A school's environment score is based on attendance rates and responses from surveys given to teachers, parents, and students. The other two scores are based on student performance in state math and English Language Arts (ELA) examinations. While student performance measures rely on level scores, student progress measures rely on growth or changes in student scores over years. The program attaches consequences to the letter grades. Higher grade (A) schools are eligible for increases in per-pupil funding and bonuses for principals. Schools receiving "F" or "D" grades are required to implement "school improvement measures and target setting." Low-performing (F- and D-schools) are also threatened with changes in their principal, and possible restructuring and closure if they continue to receive poor grades. The program also makes students in F-schools eligible to transfer to better performing schools.

> *As in Florida's voucher program, public schools in New York face valid sanctions if they fail to perform. Therefore, incentives faced by New York's low-performing schools are similar to those faced by the F-schools in Florida, and one would expect a similar response from them.*

Although the Progress Report program does not have a voucher element, it is in many ways similar to the Florida voucher program; indeed, its design was based on the Florida program. Like the Florida program, it embeds sanctions in an accountability framework with consequences/sanctions imposed on low-performing schools if they fail to improve. Additionally, the criteria of the New York City program that make students in low-performing schools eligible to transfer to other higher performing schools are similar to those of Florida's program. The only distinction is that in New York, students can transfer to public schools only—not to private schools, as in the Florida program. The threat of removal of the principal and the possibility of restructuring are sanctions imposed over and above the transfer option. These sanctions are credible and pose a valid threat to administrators. For example, as reported in Rockoff (2008), "Seven schools receiving an F and two schools receiving a D were told in December of 2007 that they would be closed immediately or phased out after the school year 2007-08. . . . Additionally, 17 percent of the remaining F-school principals (and 12 percent of the D-school principals) did not return in the school year 2008-09, relative to 9 percent of principals receiving a C, B, or A grade."

Thus, as in Florida's voucher program, public schools in New York face valid sanctions if they fail to perform. Therefore, incentives faced by New York's low-performing schools are similar to those faced by the F-schools in Florida, and one would expect a similar response from them. Accordingly, the above analysis would indicate that low-performing schools under the Progress Report program would have an incentive to improve. In fact, there is some evidence in favor of such improvement. In 2009, 82 percent of students passed in math and 69 percent in English, up from 42 percent and 38 percent, respectively, in 2002. Earlier, all five boroughs of New York City ranked toward the bottom in the state; now Queens and Staten Island rank toward the top in elementary-school math scores. The racial achievement gap in passing rates has been closed by half in some tests. (Statistics are from Elissa Gootman and Robert Gebeloff, *New York Times*, August 4, 2009.) Gootman and Gebeloff also report:

> At Public School 398 in Brownsville, Brooklyn, 77 percent of students passed the math tests this year and 60 percent passed English, up from 56 and 43 percent last year. Gene McCarthy, a fifth-grade teacher, attributed the improvement to "a tremendous amount of test prep," but said that with a little creativity on his part, "ultimately I think it's learning." The principal, Diane Danay-Caban, said at P.S. 398, which had struggled for years with low scores and discipline problems, she has come to feel that the push to raise scores has brought genuine gains in knowledge.

Rockoff and Turner (2008) find that schools labeled "F" improved their performance in both ELA and math, with larger effects in math. Winters (2008), analyzing the same program, finds improvement of F-schools in math, although he finds no such effect in ELA.

NCLB, a major reform of the Elementary and Secondary Education Act, was signed into law on January 8, 2002. The states, including New York, implemented it soon thereafter. In compliance with the law, New York established Adequate Yearly Progress (AYP) targets, and all schools were graded on the basis of the targets. AYP is determined based on each school's progress toward meeting the state proficiency level for all students in English language arts, mathematics, science, as well as the high-school graduation rate. Schools are held accountable for the achievement of students of different races and ethnic groups, students with disabilities, students with limited English proficiency, and students of low-income families. Schools must also have an average over two years of 95 percent of their students participating in state tests. If a school does not meet requirements in any one of these categories, it is said to miss AYP. Schools that receive Title I money are subject to NCLB sanctions if they miss AYP in two consecutive years. A school missing AYP for two consecutive years is required to provide public school choice to students. A school missing AYP for three consecutive years is required to provide supplemental educational services (such as tutoring) in addition to the above sanctions. Missing AYP for four consecutive years leads to corrective action in addition to the above sanctions; for five consecutive years, it results in restructuring in addition to the above sanctions. Thus, sanctions start with two years of missed AYP and escalate from there.

Only a fraction of eligible students took advantage of the transfer option in New York as well as in the nation as a whole. This result is attributable mainly to two factors: the absence of an adequate number of spaces in nearby schools and the lack of adequate information.

While NCLB does not have any voucher component, the accountability-sanctions component is similar in spirit to that of Florida's voucher program. In fact, the design of NCLB was based on that program. As in the Florida program, NCLB first threatens failing schools with sanctions, and sanctions are introduced only if the schools fail to meet the predesignated targets in the following years.[21] Therefore, one would expect similar incentives to be created by NCLB and threatened schools to respond in a way similar to the F-schools under the Florida program. In other words, one would expect schools threatened by the NCLB sanctions to improve their performance in an effort to make AYP. However, it should be emphasized that these incentives and responses would be applicable only if the sanctions are credible and pose a valid threat to the affected schools. Under NCLB, though, implementation of the sanctions has been largely limited. For example, only a fraction of eligible students took advantage of the transfer option in New York as well as in the nation as a whole. This result is attributable mainly to two factors: the absence of an adequate number of spaces in nearby schools and the lack of adequate information. For example, as reported in the New York *Daily News*, "Some parents of kids in failing schools told the *Daily News* they weren't even aware they could transfer out, and some were turned away from better schools that are already overcrowded" (February 3, 2008).

The challenge to policymakers in [accountability] programs is to establish—and enforce—credible sanctions that function as valid threats to the agents (here, public schools).

In summary, both New York City's Progress Report program and NCLB have the potential to induce improvement from threatened schools, but the incentives and response ultimately depend on how functional and credible the threats under consideration are. The challenge to policymakers in such programs is to establish—and enforce—credible sanctions that function as valid threats to the agents (here, public schools). Only in such cases would the agents have an incentive to respond in the direction intended or deemed appropriate by the policymakers.

9. Conclusion

This article examines the role of program design in the context of two educational interventions in the United States—the Florida and Milwaukee school voucher programs. Even though both programs involve vouchers, their designs are quite different: the Milwaukee program makes low-income Milwaukee public school students eligible for vouchers, while the Florida system ties vouchers to low school performance. Specifically, Florida students become eligible for vouchers if

[21] Note, though, that while under NCLB all low-performing schools face stigma (embarrassment) due to public reporting of scores and grades, only Title I schools (schools that receive Title I money) are subject to sanctions.

and only if their school receives two "F" grades in a period of four years. This study shows that program design matters; indeed, the design differences have had very different incentive and performance effects on schools subject to the two programs. Specifically, the Florida program led to considerably larger improvements from the threatened schools compared with corresponding schools under the Milwaukee program. These findings are robust to several sensitivity checks.

The lessons drawn from our analysis are applicable to some of New York City's educational policies. These policies include the No Child Left Behind Act, as implemented by the state, and New York City's "Progress Report" policy. While neither of these programs has voucher components, both are accountability programs that have consequences for schools that fail to perform. In that sense, one would expect the incentives and responses generated by these programs to be similar to those created by the Florida program. Hence, the threatened schools could be expected to improve in an effort to avoid the sanctions. In fact, there is some evidence of such improvement in the affected schools, especially in schools treated by New York City's Progress Report program. However, the extent of the responses and the performance effects ultimately depends on the credibility of the sanctions and the validity of the threat posed to the affected schools.

References

Chakrabarti, R. 2007. "Vouchers, Public School Response, and the Role of Incentives: Evidence from Florida." Federal Reserve Bank of New York Staff Reports, no. 306, October.

———. 2008a. "Can Increasing Private School Participation and Monetary Loss in a Voucher Program Affect Public School Performance? Evidence from Milwaukee." Journal of Public Economics 92, no. 5-6 (June): 1371-93.

———. 2008b. "Impact of Voucher Design on Public School Performance: Evidence from Florida and Milwaukee Voucher Programs." Federal Reserve Bank of New York Staff Reports, no. 315, January.

———. 2009. "Do Vouchers Lead to Sorting under Random Private-School Selection? Evidence from the Milwaukee Voucher Program." Federal Reserve Bank of New York Staff Reports, no. 379, July.

Epple, D., and R. E. Romano. 1998. "Competition between Private and Public Schools, Vouchers, and Peer Group Effects." American Economic Review 88, no. 1 (March): 33-62.

———. 2002. "Educational Vouchers and Cream Skimming." International Economics Review 49, no. 4 (November): 1395-1435.

Figlio, D. N., and C. E. Rouse. 2006. "Do Accountability and Voucher Threats Improve Low-Performing Schools?" Journal of Public Economics 90, no. 1-2 (January): 239-55.

Greene, J. P. 2001. "An Evaluation of the Florida A-Plus Accountability and School Choice Program." Manhattan Institute for Policy Research Civic Report (with Florida State University and the Program on Education Policy and Governance at Harvard University), February.

Greene, J. P., and M. A. Winters. 2003. "When Schools Compete: The Effects of Vouchers on Florida Public School Achievement." Manhattan Institute for Policy Research Education Working Paper no. 2, August.

Hoxby, C. M. 2003a. "School Choice and School Competition: Evidence from the United States." Swedish Economic Policy Review 10, no. 2: 9-65.

———. 2003b. "School Choice and School Productivity: Could School Choice Be a Tide that Lifts All Boats?" In C. M. Hoxby, ed., The Economics of School Choice, 287-342. National Bureau of Economic Research Conference Report. Chicago: University of Chicago Press.

Hsieh, C., and M. Urquiola. 2006. "The Effects of Generalized School Choice on Achievement and Stratification: Evidence from Chile's Voucher Program." Journal of Public Economics 90, no. 8-9 (September): 1477-1503.

Innerst, C. 2000. "Competing to Win: How Florida's A+ Plan Has Triggered Public School Reform." Available at http://www.edreform.com/published_pdf/Competing_To_Win_How_Floridas_A_Plus_Plan_Has_Triggered_Public_School_Reform.pdf.

Manski, C. F. 1992. "Educational Choice (Vouchers) and Social Mobility." Economics of Education Review 11, no. 4 (December): 351-69.

McMillan, R. 2004. "Competition, Incentives, and Public School Productivity." Journal of Public Economics 88, no. 9-10 (August): 1871-92.

Nechyba, T. J. 1996. "Public School Finance in a General Equilibrium Tiebout World: Equalization Programs, Peer Effects, and Private School Vouchers." National Bureau of Economic Research Working Paper no. 5642, June.

———. 1999. "School-Finance-Induced Migration and Stratification Patterns: The Impact of Private School Vouchers." Journal of Public Economic Theory 1, no. 1 (January): 5-50.

———. 2000. "Mobility, Targeting, and Private-School Vouchers." American Economic Review 90, no. 1 (March): 130-46.

———. 2003. "Introducing School Choice into Multidistrict Public School Systems." In C. M. Hoxby, ed., The Economics of School Choice, 145-94. National Bureau of Economic Research Conference Report. Chicago: University of Chicago Press.

Rockoff, J. E., and L. J. Turner. 2008. "Short-Run Impacts of Accountability on School Quality." National Bureau of Economic Research Working Paper no. 14564, December.

West, M. R., and P. E. Peterson. 2005. "The Efficacy of Choice Threats within School Accountability Systems: Results from Legislatively Induced Experiments." Harvard University Program on Education Policy and Governance, PEPG no. 05-01.

Winters, M. A. 2008. "Grading New York: An Evaluation of New York City's Progress Report Program." Manhattan Institute for Policy Research Civic Report no. 55, November.

Witte, J. F. 2000. The Market Approach to Education: An Analysis of America's First Voucher Program. Princeton, N.J.: Princeton University Press.

Argia M. Sbordone, Andrea Tambalotti, Krishna Rao, and Kieran Walsh

Policy Analysis Using DSGE Models: An Introduction

- Dynamic stochastic general equilibrium models are playing an important role in the formulation and communication of monetary policy at many of the world's central banks.

- These models, which emphasize the dependence of current choices on expected future outcomes, have moved from academic circles to the policymaking community—but they are not well known to the general public.

- This study adds to the understanding of the DSGE framework by using a small-scale model to show how to address specific monetary policy questions; the authors focus on the causes of the sudden pickup in inflation in the first half of 2004.

- An important lesson derived from the exercise is that the management of expectations can be a more effective tool for stabilizing inflation than actual movements in the policy rate; this result is consistent with the increasing focus on central bankers' pronouncements of their future actions.

1. Introduction

In recent years, there has been a significant evolution in the formulation and communication of monetary policy at a number of central banks around the world. Many of these banks now present their economic outlook and policy strategies to the public in a more formal way, a process accompanied by the introduction of modern analytical tools and advanced econometric methods in forecasting and policy simulations. Official publications by central banks that formally adopt a monetary policy strategy of inflation targeting—such as the Inflation Report issued by the Bank of England and the monetary policy reports issued by the Riksbank and Norges Bank—have progressively introduced into the policy process the language and methodologies developed in the modern dynamic macroeconomic literature.[1]

The development of medium-scale DSGE (dynamic stochastic general equilibrium) models has played a key role in this process.[2] These models are built on microeconomic foundations and emphasize agents' intertemporal choice. The dependence of current choices on future uncertain

[1] The Bank of England has published a quarterly Inflation Report since 1993. The report sets out the detailed economic analysis and inflation projections on which the Bank's Monetary Policy Committee bases its interest rate decisions. The Riksbank and Norges Bank each publish monetary policy reports three times a year. These reports contain forecasts for the economy and an assessment of the interest rate outlook for the medium term.

[2] A simple exposition of this class of models can be found in Galí and Gertler (2007). Woodford (2003) provides an exhaustive textbook treatment.

Argia M. Sbordone is an assistant vice president and Andrea Tambalotti a senior economist at the Federal Reserve Bank of New York; Krishna Rao is a graduate student at Stanford University; Kieran Walsh is a graduate student at Yale University.
Correspondence: argia.sbordone@ny.frb.org, andrea.tambalotti@ny.frb.org

The authors thank Ariel Zetlin-Jones for his contribution to the early stages of this work. 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.

outcomes makes the models dynamic and assigns a central role to agents' expectations in the determination of current macroeconomic outcomes. In addition, the models' general equilibrium nature captures the interaction between policy actions and agents' behavior. Furthermore, a more detailed specification of the stochastic shocks that give rise to economic fluctuations allows one to trace more clearly the shocks' transmission to the economy.

The use of DSGE models as a potential tool for policy analysis has contributed to their diffusion from academic to policymaking circles. However, the models remain less well-known to the general public. To broaden the understanding of these models, this article offers a simple illustration of how an estimated model in this class can be used to answer specific monetary policy questions. To that end, we introduce the structure of DSGE models by presenting a simple model, meant to flesh out their distinctive features. Before proceeding to a formal description of the optimization problems solved by firms and consumers, we use a simple diagram to illustrate the interactions among the main agents in the economy. With the theoretical structure in place, we discuss the features of the estimated model and the extent to which it approximates the volatility and comovement of economic time series. We also discuss important outcomes of the estimation—namely, the possibility of recovering the structural shocks that drive economic fluctuations as well as the historical behavior of variables that are relevant for policy but are not directly observable. We conclude by applying the DSGE tool to study the role of monetary policy in a recent episode of an increase in inflation. The lesson we emphasize is that, while they are a very stylized representation of the real economy, DSGE models provide a disciplined way of thinking about the economic outlook and its interaction with policy.[3]

This article offers a simple illustration of how an estimated [DSGE] model . . . can be used to answer specific monetary policy questions.

We work with a small model in order to make the transmission mechanism of monetary policy, whose basic contours our model shares with most DSGE specifications, as transparent as possible. Therefore, the model focuses on the behavior of only three major macroeconomic variables: inflation, GDP growth, and the short-term interest rate. However, the basic framework that we present could easily be enriched to provide more details on the structure of the economy. In fact, a key advantage of DSGE models is that they share core assumptions on the behavior of households and firms, which makes them easily scalable to include details that are relevant to address the question at hand. Indeed, several extensions of the basic framework presented here have been developed in the literature, including the introduction of wage stickiness and frictions in the capital accumulation process (see the popular model of Smets and Wouters [2007]) and a treatment of wage bargaining and labor market search (Gertler, Sala, and Trigari 2008).[4] Recently, the 2008 financial crisis has highlighted one key area where DSGE models must develop: the inclusion of a more sophisticated financial intermediation sector. There is a large body of work under way to model financial frictions within the baseline DSGE framework—work that is very promising for the study of financial intermediation as a source and conduit of shocks as well as for its implications for monetary policy. However, this last generation of models has not yet been subjected to extensive empirical analysis.

A key advantage of DSGE models is that they share core assumptions on the behavior of households and firms, which makes them easily scalable to include details that are relevant to address the question at hand.

Our study is organized as follows. Section 2 describes the general structure of our model while Section 3 illustrates its construction from microeconomic foundations. Section 4 briefly describes our approach to estimation and presents some of the model's empirical properties. In Section 5, we use the model to analyze the inflationary episode of the first half of 2004. Section 6 concludes.

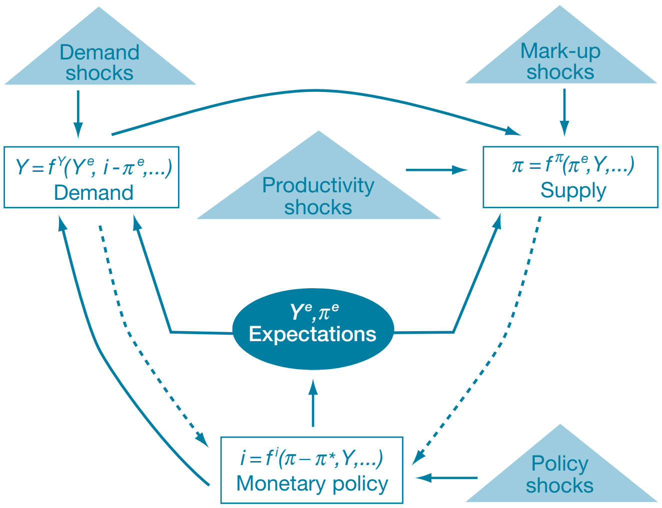
2. DSGE Models and Their Basic Structure

Dynamic stochastic general equilibrium models used for policy analysis share a fairly simple structure, built around three interrelated blocks: a demand block, a supply block, and a

[3]Adolfson et al. (2007) offer a more extended illustration of how DSGE models can be used to address questions that policymakers confront in practice. Erceg, Guerrieri, and Gust (2006) illustrate policy simulations with an open-economy DSGE model.

[4] Some of these larger DSGE models inform policy analysis at central banks around the world: Smets and Wouters (2007) of the European Central Bank; Edge, Kiley, and Laforte (2008) of the Federal Reserve System; and Adolfson et al. (2008) of the Riksbank.

The Basic Structure of DSGE Models



monetary policy equation. Formally, the equations that define these blocks derive from microfoundations: explicit assumptions about the behavior of the main economic actors in the economy—households, firms, and the government. These agents interact in markets that clear every period, which leads to the "general equilibrium" feature of the models. Section 3 presents the microfoundations of a simple DSGE model and derives the equations that define its equilibrium. But first, we begin by introducing the basic components common to most DSGE models with the aid of a diagram.

In the diagram, the three interrelated blocks are depicted as rectangles. The demand block determines real activity (Y) as a function of the ex ante real interest rate—the nominal rate minus expected inflation ($i - \pi^e$)—and of expectations about future real activity (Y^e). This block captures the idea that, when real interest rates are temporarily high, people and firms would rather save than consume or invest. At the same time, people are willing to spend more when future prospects are promising (Y^e is high), regardless of the level of interest rates.

The line connecting the demand block to the supply block shows that the level of activity (Y) emerging from the demand block is a key input in the determination of inflation (π), together with expectations of future inflation (π^e). In prosperous times, when the level of activity is high, firms must increase wages to induce employees to work longer hours. Higher wages increase marginal costs, putting pressure on prices and generating inflation. Moreover, the higher inflation is expected to be in the future, the higher is this increase in prices, thus contributing to a rise in inflation today.

The determination of output and inflation from the demand and supply blocks feeds into the monetary policy block, as indicated by the dashed lines. The equation in that block describes how the central bank sets the nominal interest rate, usually as a function of inflation and real activity. This reflects the tendency of central banks to raise the short-term interest rate when the economy is overheating as well as when inflation rises and to lower it in the presence of economic slack. By adjusting the nominal interest rate, monetary policy in turn affects real activity and through it inflation, as represented by the line flowing from the monetary policy block to the demand block and then to the supply block. The policy rule therefore closes the circle, giving us a complete model of the relationship between three key endogenous variables: output (Y), inflation (π), and the nominal interest rate (i).

While this brief description appears static, one of the fundamental features of DSGE models is the dynamic interaction between the blocks—hence, the "dynamic" aspect of the DSGE label—in the sense that expectations about the future are a crucial determinant of today's outcomes. These expectations are pinned down by the same mechanism that generates outcomes today. Therefore, output and inflation tomorrow, and thus their expectations as of today, depend on monetary policy tomorrow in the same way as they do today—of course, taking into account what will happen from then on into the infinite future.

One of the fundamental features of DSGE models is the dynamic interaction between the [three interrelated] blocks—hence, the "dynamic" aspect of the DSGE label—in the sense that expectations about the future are a crucial determinant of today's outcomes.

The diagram highlights the role of expectations and the dynamic connections between the blocks that they create. The influence of expectations on the economy is represented by the arrows, which flow from monetary policy to the demand and then the supply block, where output and inflation are determined. This is to emphasize that the conduct of monetary policy has a large influence on the formation of expectations. In fact, in DSGE models, expectations are the main channel through which policy affects the economy, a feature that is consistent with the close attention paid by financial markets and the public to the pronouncements of central banks on their likely course of action.

The last component of DSGE models captured in the diagram is their stochastic nature. Every period, random exogenous events perturb the equilibrium conditions in each block, injecting uncertainty in the evolution of the economy and thus generating economic fluctuations. Without these shocks, the economy would evolve along a perfectly predictable path, with neither booms nor recessions. We represent these shocks as triangles, with arrows pointing toward the equilibrium conditions on which they directly impinge. Markup and productivity shocks, for example, affect the pricing and production decisions of firms that underlie the supply block, while demand shocks capture changes in the willingness of households to purchase the goods produced by those firms.

3. Microfoundations of a Simple DSGE Model

We present the microfoundations of a small DSGE model that is simple enough to fit closely into the stylized structure outlined in our diagram. Our objective is to describe the basic components of DSGE models from a more formal perspective, using the mathematical language of economists, but avoiding unnecessary technical details. Despite its simplicity, our model is rich enough to provide a satisfactory empirical account of the evolution of output, inflation, and the nominal interest rate in the United States in the last twenty years, as we discuss in the next section.

Given the constraints we impose on this treatment for the sake of simplicity, our model lacks many features that are standard in the DSGE models that central banks typically use. For example, we ignore the process of capital accumulation,

> *Despite its simplicity, our model is rich enough to provide a satisfactory empirical account of the evolution of output, inflation, and the nominal interest rate in the United States in the last twenty years.*

which would add another dimension—investment decisions by firms—to the economy's demand block. Nor do we attempt to model the labor market in detail: for example, we make no distinction between the number of hours worked by each employee and the number of people at work, an issue that is hard to overlook in a period with unemployment close to 10 percent. Finally, we exclude any impediment to the smooth functioning of financial markets and assume that the central bank can perfectly control the short-term interest rate—the only relevant rate of return in the economy. The 2008 financial crisis has proved that this set of assumptions can fail miserably in some circumstances and has highlighted the need for a more nuanced view of financial markets within the current generation of DSGE models, as we observe in our introduction.

3.1. The Model Economy

Our model economy is populated by four classes of agents: a representative household, a representative final-good-producing firm (*f*-firm), a continuum of intermediate firms (*i*-firms) indexed by $i \in [0, 1]$, and a monetary authority. The household consumes the final good and works for the *i*-firms. Each of these firms is a monopolist in the production of a particular intermediate good i, for which it is thus able to

> *Our model economy is populated by four classes of agents: a representative household, a representative final-good-producing firm, . . . a continuum of intermediate firms, . . . and a monetary authority.*

set the price. The *f*-firm packages the differentiated goods produced by the *i*-firms and sells the product to households in a competitive market. The monetary authority sets the nominal interest rate.

The remainder of this section describes the problem faced by each economic agent, shows the corresponding optimization conditions, and interprets the shocks that perturb these conditions. These optimization conditions result in dynamic relationships among macroeconomic variables that define the three model blocks described above. Together with market clearing conditions, these relationships completely characterize the equilibrium behavior of the model economy.

3.2. Households and the Aggregate Demand Block

At the core of the demand side of virtually all DSGE models is a negative relationship between the real interest rate and desired spending. In our simple model, the only source of spending is consumption. Therefore, the negative relationship between the interest rate and demand emerges from the consumption decision of households.

We model this decision as stemming from the optimal choice of a very large representative household—the entire U.S. population—which maximizes its expected discounted lifetime utility, looking forward from an arbitrary date t_0

$$\underset{\{B_{t_0+s}, C_{t_0+s}, [H_{t_0+s}(i)]_{i\in[0,1]}\}_{s=0}^{\infty}}{Max} E_{t_0}\sum_{s=0}^{\infty}\beta^s\Big\{b_{t_0+s}\Big[\log(C_{t_0+s} -\eta C_{t_0+s-1})-\int_0^1 v(H_{t_0+s}(i))di\Big]\Big\}$$

subject to the sequence of budget constraints

$$P_t C_t + \frac{B_t}{R_t} \le B_{t-1} + \int_0^1 w_t(i)H_t(i)di,$$

for $t = t_0, t_0+1, \ldots, \infty$, and given B_{t_0-1}. The members of this household, we call them "Americans," like consumption, C_t, but dislike the number of hours they spend at work, H_t, to an extent described by the convex function v. The utility flow from consumption depends on current as well as past consumption, with a coefficient η. As a result of this "habit," consumers are unhappy if their current consumption is low, but also if it falls much below the level of their consumption in the recent past. To afford consumption, Americans work a certain amount of hours $H_t(i)$ in each of the i-firms, where they earn an hourly nominal wage $W_t(i)$ which they take as given when deciding how much to work.[5] With the income thus earned, they can purchase the final good at price P_t or save by accumulating one-period discount government bonds B_t, whose gross rate of return between t and $t+1$ is R_t.

From the perspective of time t, the household discounts utility in period $t+1$ by a time-varying factor $\beta b_{t+1}/b_t$, where b_{t+1}/b_t is an exogenous stochastic process. Changes in b_{t+1}/b_t represent shocks to the household's impatience. When b_{t+1} increases relative to b_t, for example, the household cares more about the future and thus wishes to save more and consume less today, everything else equal. In this respect, b_{t+1}/b_t acts as a traditional *demand shock*, which affects desired consumption and saving exogenously. A persistent increase in b_{t+1}/b_t is one way of interpreting the current macroeconomic situation in the United States, in which households have curtailed their consumption—partly to build their savings. Of course, in reality there are many complex reasons behind this observed change in behavior, and an increase in people's concern about the future is surely one of them. For simplicity, the model focuses on this one reason exclusively.

[5] In equilibrium, the wage—and thus the number of hours worked—will settle at the level at which the supply of labor by the household equals the demand of labor by firms. This labor demand in turn is determined by the need of firms to hire enough workers to satisfy the demand for their products, as we describe in the next section.

To find the solution to the optimal problem above, we form the Lagrangian

$$L = E_{t_0}\sum_{s=0}^{\infty}\Big\{\beta^s\Big[b_{t_0+s}\Big(\log(C_{t_0+s}-\eta C_{t_0+s-1}) - \int_0^1 v(H_{t_0+s}(i))di\Big) - \Lambda_{t_0+s}(P_{t_0+s}C_{t_0+s} + B_{t_0+s}R^{-1}_{t_0+s}) - B_{t_0+s-1} - \int_0^1 W_{t_0+s}(i)H_{t_0+s}(i)di\Big)\Big]\Big\},$$

with first-order conditions

(3.1a) $$\frac{\partial L}{\partial B_t} : \Lambda_t = \beta E_t[\Lambda_{t+1}]R_t$$

(3.1b) $$\frac{\partial L}{\partial C_t} : \frac{\Lambda_t}{b_t}P_t = \frac{1}{C_t - \eta C_{t-1}} - \eta E_t\Big[\frac{\beta b_{t+1}/b_t}{C_{t+1}-\eta C_t}\Big].$$

for $t = t_0, t_{0+1}, \ldots \infty$ and

(3.2) $$\frac{\partial L}{\partial H_t(i)} : \frac{v'(H_t(i))}{\Lambda_t/b_t} = W_t(i)$$

for $t = t_0, t_{0+1}, \ldots \infty$ and $\forall i \in [0,1]$, together with the sequence of budget constraints.

These conditions yield a fully state-contingent plan for the household's choice variables—how much to work, consume, and save in the form of bonds—looking forward from the planning date t_0 and into the foreseeable future. At any point in time, the household is obviously uncertain about the way in which this future will unfold. However, we assume that the household is aware of the kind of random external events, or shocks, that might affect its decisions and, crucially, that it knows the probability with which these shocks might occur. Therefore, the household can form expectations about future outcomes, which are one of the inputs in its current choices. We assume that these expectations are rational, meaning that they are based on the same knowledge of the economy and of the shocks that buffet it as that of the economist constructing the model. We use the notation $E_t[X_{t+s}]$ to denote expectations formed at time t of any future variable X_{t+s}, as in equations 3.1, for example. The optimal plan, then, is a series of instructions on how to behave in response to the realization of each shock, given expectations about the future, rather than a one-time decision on exactly how much to work, consume, and save on each future date.

Together, the optimality conditions in equations 3.1 establish the negative relationship between the interest rate and desired consumption that defines the demand side of the model. The nature of this relationship is more transparent in the special case of no habit in consumption ($\eta = 0$), when we can combine the two equations to obtain

$$(3.3)\qquad \frac{1}{C_t} = E_t\left[\frac{\beta b_{t+1}}{b_t}\frac{1}{C_{t+1}}\frac{R_t}{P_{t+1}/P_t}\right].$$

According to this so-called Euler equation, desired consumption decreases when the (gross) real interest rate $\left(\frac{R_t}{P_{t+1}/P_t}\right)$ increases, when expected future consumption decreases, and when households become more patient (b_{t+1} rises).

A log-linear approximation of the Euler equation (3.3), after some manipulation, gives

$$(3.4)\qquad y_t = E_t y_{t+1} - (i_t - E_t \pi_{t+1}) - \delta_t ,$$

where $\pi_t \equiv \log P_t / P_{t-1}$ is the quarterly inflation rate, $i_t \equiv \log R_t$ is the continuously compounded nominal interest rate, $\delta_t \equiv E_t \log(\beta b_{t+1} / b_t)$ is a transformation of the demand shock, and $y_t \equiv \log Y_t$ is the logarithm of total output. In this expression, we can substitute consumption of the final good C_t with its output Y_t because in our model consumption is the only source of demand for the final good. Therefore, market clearing implies $Y_t = C_t$.

In this framework, equation 3.4 is similar to a traditional IS equation, since it describes the relationship between aggregate activity y_t and the ex ante real interest rate $i_t - E_t \pi_{t+1}$, which must hold for the final-good market to clear. Unlike a traditional IS relationship, though, this equation is dynamic and forward looking, as it involves current and future expected variables. In particular, it establishes a link between current output and the entire future expected path of real interest rates, as we see by solving the equation forward

$$(3.5)\qquad y_t = -E_t \sum_{s=0}^{\infty} (i_{t+s} - \pi_{t+s+1} - \delta_{t+s}).$$

Through this channel, expectations of future monetary policy directly affect current economic conditions. In fact, this equation shows that future interest rates are just as important to determine today's output as the current level of the short-term rate, as we describe in our discussion of the role of policy expectations.

In our full model, the Euler equation is somewhat more complicated than in equation 3.4 because of the consumption habit ($\eta \neq 0$), which is a source of richer, and more realistic, output dynamics in response to changes in the interest rate. Nevertheless, these more intricate dynamics do not change the qualitative nature of the relationship between real rates and demand.

The third first-order condition of the household optimization problem, equation 3.2, represents the labor supply decision. It says that Americans are willing to work more hours in firms that pay a higher wage and at times when wages are higher, at least for differences in wages modest enough to have no significant effect on their income.[6] Large wage changes, in fact, would trigger an income effect and lead the now richer workers to curtail their labor supply. Mathematically, workers with higher income could afford more consumption, which would lead to a drop in the marginal utility Λ_t and thus to a decrease in labor supply at any given wage level.

We can think of the labor supply schedule (equation 3.2) as a relationship determining the wage that firms must pay to induce Americans to work a certain number of hours. In prosperous times, when demand is high and firms are producing much, firms require their labor force to work long hours and they must correspondingly pay higher hourly wages. This is an important consideration in the production and pricing decisions of firms, as we discuss in the next section.

3.3. Firms and the Aggregate Supply Block

The supply block of a DSGE model describes how firms set their prices as a function of the level of demand they face. Recall that in prosperous times, demand is high and firms must pay their workers higher wages. As a result, their costs increase as do their prices. In the aggregate, this generates a positive relationship between inflation and real activity.

In terms of microfoundations, establishing this relationship requires some work, since firms must have some monopoly power to set prices. This is why our production structure includes a set of monopolistic *i*-firms, which set prices, as well as an *f*-firm, which simply aggregates the output of the *i*-firm into the final consumption good. Because all the pricing action occurs within the *i*-firms, we focus on their problem and omit that of the *f*-firm.

Intermediate firm i hires $H_t(i)$ units of labor of type i on a competitive market to produce $Y_t(i)$ units of intermediate good i with the technology

$$(3.6)\qquad Y_t(i) = A_t H_t(i) ,$$

where A_t represents the overall efficiency of the production process. We assume that A_t follows an exogenous stochastic process, whose random fluctuations over time capture the unexpected changes in productivity often experienced by modern economies—for example, the productivity boom in the mid-1990s that followed the mass adoption of information technology. We call this process an aggregate *productivity shock*, since it is common to all firms.

[6] Labor supply is upward sloping because ν' is an increasing function, as ν is convex. In other words, people dislike working an extra hour more intensely when they are already working a lot rather than when they are working little.

The market for intermediate goods is monopolistically competitive, as in Dixit and Stiglitz (1977), so firms set prices subject to the requirement that they satisfy the demand for their good. This demand comes from the *f*-firm and takes the form

$$(3.7) \qquad Y_t(i) = Y_t\left(\frac{P_t(i)}{P_t}\right)^{-\theta_t},$$

where $P_t(i)$ is the price of good i and θ_t is the elasticity of demand. When the relative price of good i increases, its demand falls relative to aggregate demand by an amount that depends on θ_t.

Moreover, we assume that firms change their prices only infrequently. The fact that firms do not adjust prices continuously, but leave them unchanged in some cases for long periods of time, should be familiar from everyday experience and is well established in the economic literature (for example, Bils and Klenow [2004]; Nakamura and Steinsson [2008]). To model this fact, we follow Calvo (1983) and assume that in every period only a fraction $1-\alpha$ of firms is free to reset its price while the remaining fraction maintains its old price.[7]

The subset of firms that are able to set an optimal price at t, call it $\Omega_t \subset [0,1]$, maximize the discounted stream of expected future profits, taking into account that s periods from now there is a probability α^s that they will be forced to retain the price chosen today. The objective function of each of these firms is therefore

$$\underset{P_t(i)}{\text{Max}}\ E_t\sum_{s=0}^{\infty}\alpha^s\frac{\beta^s\Lambda_{t+s}}{\Lambda_t}\{P_t(i)Y_{t+s}(i)-W_{t+s}(i)H_{t+s}(i)\}$$

for all $i \in \Omega_t$, subject to the production function 3.6 and to the additional constraint that they must satisfy the demand for their product at every point in time

$$(3.8) \qquad Y_{t+s}(i) = Y_{t+s}\left(\frac{P_t(i)}{P_{t+s}}\right)^{-\theta_{t+s}},$$

for $s = 0,1,\ldots,\infty$. Profits, which are given by total revenue at the price chosen today, $P_t(i)\,Y_{t+s}(i)$ minus total costs $W_{t+s}(i)H_{t+s}(i)$, are discounted by the multiplier $\beta^s\Lambda_{t+s}/\Lambda_t$, sometimes called a stochastic discount factor, which translates dollar profits in the future into a current dollar value.

The first-order condition of this optimization problem is

$$(3.9) \quad E_t\sum_{s=0}^{\infty}(\alpha\beta)^s\Lambda_{t+s}Y_{t+s}P_{t+s}^{\theta_{t+s}-1}\left[P_t^*(i)-\mu_{t+s}\frac{W_{t+s}(i)}{A_{t+s}}\right]=0,$$

for all $i \in \Omega_t$, where $P_t^*(i)$ denotes the optimal price chosen by firm i, $W_{t+s}(i)/A_{t+s}$ is the firm's nominal marginal cost at time $t+s$, and $\mu_{t+s} = \frac{\theta_{t+s}-1}{\theta_{t+s}}$ is its desired mark-up—the mark-up it would charge if prices were flexible. As rational monopolists, optimizing firms set their price as a mark-up over their marginal cost. However, this relationship holds in expected present discounted value, rather than every period, since a price chosen at time t will still be in effect with probability α^s in period $t+s$.

We can rewrite the marginal cost of a firm that at time $t+s$ is still forced to retain the price $P_t(i)$ as

$$(3.10) \qquad S_{t+s}(i) \equiv \frac{W_{t+s}(i)}{A_{t+s}} = \frac{v'[H_{t+s}(i)]}{\Lambda_{t+s}/b_{t+s}}\frac{1}{A_{t+s}}$$

$$= \frac{v'\left(\frac{Y_{t+s}}{A_{t+s}}\left(\frac{P_t(i)}{P_{t+s}}\right)^{-\theta_{t+s}}\right)}{A_{t+s}\Lambda_{t+s}/b_{t+s}},$$

where we use the labor supply relation 3.2 to substitute for the wage as well as the production function 3.6 and the demand function 3.8 to give us an expression for the labor demand $H_{t+s}(i)$.[8] Inserting this expression into the first-order condition 3.9, we see that the solution to the optimal pricing problem is the same for all firms in the set Ω_t, since it depends only on the aggregate variables $\{Y_{t+s}, A_{t+s}, P_{t+s}, \Lambda_{t+s}\}_{s=0}^{\infty}$. We denote this common optimal price as P_t^*.

The equation for the desired mark-up—$\mu_{t+s} = \frac{\theta_{t+s}-1}{\theta_{t+s}}$, also known as Lerner's formula—says that monopolists facing a more rigid demand optimally charge a higher mark-up, and thus higher prices, since their clients are less sensitive to changes in the latter. We assume that this sensitivity—the elasticity of demand—and thus the desired mark-up, follows an exogenous stochastic process. Positive realizations of this *desired mark-up shock*, which correspond to a fall in the elasticity of demand, represent an increase in firms' market power, to which they respond by increasing their price.

Equation 3.9, together with the definition of the aggregate price level as a function of newly set prices P_t^* and of the past price index P_{t-1}

$$P_t \equiv [(1-\alpha)P_t^{*(1-\theta_t)} + \alpha P_{t-1}^{1-\theta_t}]^{\frac{1}{1-\theta_t}}$$

yields an approximate New Keynesian Phillips curve—a relationship between current inflation, future expected inflation, and real marginal cost—of the form

$$(3.11) \qquad \pi_t = \xi s_t + \beta E_t\pi_{t+1} + u_t,$$

where $u_t = \xi \log u_t$ is a transformation of the mark-up shock and $s_t \equiv \log(S_t/P_t)$ is the logarithm of the *real* marginal cost.[9] The sensitivity of inflation to changes in the marginal cost, ξ, depends on the frequency of price adjustment α, as well as on

[7] In our estimated model in Section 5, we actually assume that the α fraction of firms that cannot choose their price freely can in fact adjust it in part to catch up with recent inflation. This assumption improves the model's ability to fit the data on inflation, but it would complicate our presentation of the model's microfoundations without altering its basic message.

[8] These substitutions are equivalent to "solving" for equilibrium in the labor market.

[9] Variables are in logarithms, because equation 3.11, like equation 3.4, is obtained by a log-linear approximation.

other structural parameters, according to $\xi \equiv \frac{(1-\alpha)(1-\alpha\beta)}{\alpha(1+\omega\theta)}$, where $\omega \equiv \frac{\nu'' H}{\nu'}$ is the elasticity of the marginal disutility of work, while θ is the average value of the elasticity of demand θ_t.

This Phillips curve, together with the expression for marginal costs (3.10), provides the relationship between inflation and real activity that defines the supply block of the model. In fact, we see from equation 3.10 that marginal cost depends on the level of aggregate activity, among other factors. Higher economic activity leads to higher wages and marginal costs. Thus, firms increase their prices, boosting aggregate inflation.

Another important feature of the Phillips curve is that it is forward looking, just as the Euler equation in the previous section is. As in that case, therefore, we can iterate equation 3.11 forward to obtain

$$\pi_t = E_t \sum_{s=0}^{\infty} \beta^s (\xi s_{t+s} + u_{t+s}),$$

which highlights how inflation today really depends on the entire future expected path of marginal costs, and through those, of real activity. But this path depends in turn on expectations about interest rates, and thus on the entire future course of monetary policy, as equation 3.5 shows. Hence, we have the crucial role of policy expectations for the determination of current economic outcomes in this model, a feature we discuss in Section 2.

Monetary Policy

Recall that when the interest rate—current and expected—is low, people demand more consumption goods (equation 3.5). But if demand is high, firms' marginal costs increase and so do their prices. The end result is inflation. The opposite is true when the interest rate is high. But where does the interest rate come from? In DSGE models, as in the real world, the short-term interest rate is set by monetary policy. In practice, this is a decision made by a committee (the Federal Reserve's Federal Open Market Committee, or FOMC) using various inputs: large data sets, projections from several models, and the judgment of policymakers. Despite the apparent complexity of the process, Taylor (1993) famously demonstrated that it can be reasonably well approximated by assuming that the Federal Reserve raises the federal funds rate when inflation and/or output is "high" with respect to some baseline. This behavior is assumed in almost all variants of DSGE models, although the definition of the appropriate baselines is somewhat controversial.

In our model, we assume that interest rates are set according to the policy rule

(3.12) $$i_t = \rho i_{t-1} + (1-\rho)[r_t^e + \pi_t^* + \phi_\pi(\pi_t^{4Q} - \pi_t^*) + \phi_y(y_t - y_t^e)] + \varepsilon_t^i,$$

where r_t^e, π_t^*, and y_t^e are the baselines for the real interest rate, inflation, and output, respectively, and $\pi_t^{4Q} \equiv (\log P_t / P_{t-4})$ is the rate of inflation over the previous four quarters. The monetary policy shock ε_t^i, a random variable with mean zero, captures any deviation of the observed nominal interest rate from the value suggested by the rule. This rule implies that, if inflation and output rise above their baseline levels, the nominal interest rate is lifted over time above its own baseline, $r_t^e + \pi_t^*$, by amounts dictated by the parameters φ_π and φ_y and at a speed that depends on the coefficient ρ. The higher policy rate, which is expected to persist even after output and inflation have returned to normal, exerts a restraining force on the economy—curbing demand, marginal costs, and inflation. In this respect, π_t^* and y_t^e can be regarded as targets of monetary policy—the levels of inflation and output that the central bank considers consistent with its mandate—and therefore do not elicit either a restrictive or a stimulative policy.

When output is at its efficient level, inflation is not stable, as policymakers would like it to be, but fluctuates because of the presence of mark-up shocks. This is the essence of the monetary policy trade-offs in the economy.

In equation 3.12, we denote the central bank's objective in terms of production as y_t^e, the "efficient" level of output. This unobserved variable can be derived from the microfoundations of the model.[10] It represents the level of output that would prevail in the economy if we could eliminate at once all distortions—namely, force the *i*-firms to behave competitively rather than as monopolists and allow them to change their prices freely. The level of activity that would result from such a situation is ideal from the perspective of the representative household in the model, as its name suggests. This is what makes it a suitable target for monetary policy. However, when output is at its efficient level, inflation is not stable, as policymakers would like it to be, but fluctuates because of the presence of mark-up shocks. This is the essence of the monetary policy trade-offs in the economy. Achieving the

[10] The precise mathematical definition of efficient output in the model is irrelevant for our purposes. We present in Section 4 an estimate of the behavior of this variable over the last twenty years.

efficient level of output requires undesirable movements in inflation. In contrast, a stable inflation implies deviations from the efficient level of output. The two objectives cannot be reconciled, but must be traded off of each other.

Related to the efficient level of output is the efficient real interest rate, r_t^e, which is the rate of return we would observe in the efficient economy described above. This definition implies that, when the actual real interest rate is at its efficient level and is expected to remain there in the future, output will also be at its efficient level. This is why we include r_t^e in our definition
of the baseline interest rate.

The other component of this baseline rate is the inflation target π_t^*. We allow this target to vary slowly over time to accommodate the fact that inflation hovered at about 4 percent for a few years around 1990 before declining to nearly 2 percent after the recession that ended in 1991. Nominal interest rates were correspondingly higher in the first period, thus implying a stable average for the real interest rate. We now present our estimate of the evolution of the inflation target.

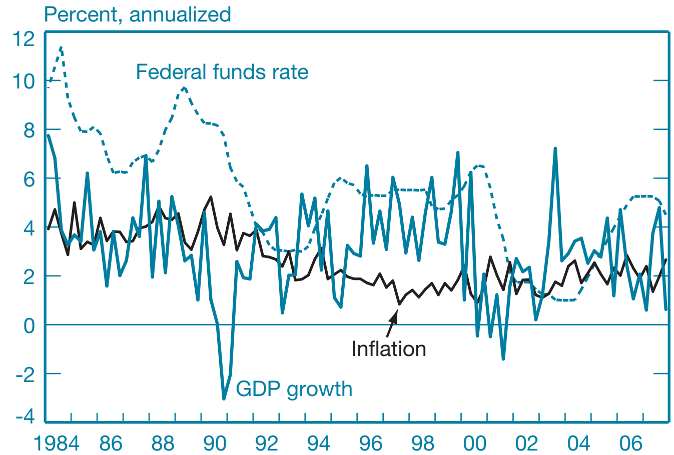
4. Empirical Approach and Estimation Results

We estimate our model using data on the growth rate of real GDP to measure output growth, Δy_t, the growth rate of the personal consumption expenditures chain price index excluding food and energy (core PCE) to measure inflation, π_t, and the quarterly average of the monthly effective federal funds rate to measure the nominal interest rate, i_t. We measure inflation by core PCE, rather than by a more comprehensive measure, because the monetary policy debate in the United States tends to focus on this index.

Our data span the period 1984:1 to 2007:4 (Chart 1). This is the longest possible data set over which it is reasonable to argue that U.S. monetary policy can be represented by a stable interest rate rule. It follows the period of extremely high interest rates in the early 1980s that brought inflation under control. However, in the first few years of this sample, inflation and the federal funds rate were still relatively high, with a fairly abrupt reduction taking place around the 1991 recession. We capture this low-frequency movement in inflation and the nominal interest rate by including the slow-moving inflation target π_t^* in the policy rule.

We use Bayesian methods to characterize the posterior distribution of the parameters of the model. This distribution combines the model's likelihood function with prior information on the parameters, using techniques surveyed, for

Chart 1
Observable Variables



Sources: Bureau of Economic Analysis; Board of Governors of the Federal Reserve System.

example, by An and Schorfheide (2007).[11] A discussion of these methods is beyond the scope of this article. Instead, we focus on the implications of these estimates for some key properties of the model. Our objective is to show that the estimated model provides a good fit to the data across many dimensions, but also to highlight some of the model's most notable shortcomings.

4.1. Moment Comparisons

We compare the second moments implied by the estimated model with those measured in the data. Table 1 presents the standard deviations of the observable variables, reported as annualized percentages. In the model column, we report the median and 5th and 95th percentiles of the standard deviations across draws from the model's posterior. This interval reflects the uncertainty on the structural parameters—and thus on the model-implied moments—encoded in the parameters' posterior distribution. In the data column, we report the median and 5th and 95th percentiles of the empirical standard deviations in the data. This interval represents the uncertainty on the true empirical moments because of the small sample available for their estimation.

Our model does a very good job replicating the volatilities in the data. It captures the standard deviation of output growth and replicates quite closely the volatility of the nominal interest rate, although it overestimates the standard deviation of

[11] The technical appendix provides information on the priors for the parameters.

Table 1
Model-Implied and Empirical Standard Deviations
Percent

Variable	Model	Data
GDP growth	2.03	2.03
	[1.79, 2.37]	[1.74, 2.27]
Core PCE inflation	1.41	1.15
	[0.98, 2.40]	[0.67, 1.38]
Federal funds rate	2.23	2.46
	[1.61, 3.56]	[1.55, 2.94]

Source: Authors' calculations.

Notes: The table reports the standard deviations of the observable variables. The model-implied standard deviations are medians across draws from the posterior; the 5th and 95th posterior percentiles across those same draws are in brackets. The empirical standard deviations are medians across bootstrap replications of a VAR(4) fit to the data; the 5th and 95th percentiles across those same replications are in brackets. PCE is personal consumption expenditures.

inflation. The ability of the model to accurately reproduce the volatility of the observable variables is not a preordained conclusion, even if we freely estimate the standard deviations of the shocks. The reason is that a likelihood-based estimator tries to match the entire autocovariance function of the data, and thus must strike a balance between matching standard deviations and all the other second moments—namely, autocorrelations and cross-correlations.

These other moments are displayed in Chart 2. The black line represents the model-implied correlation, with the shaded area representing a 90 percent posterior interval. The solid blue

Our model does a very good job replicating the volatilities in the data. It captures the standard deviation of output growth and replicates quite closely the volatility of the nominal interest rate, although it overestimates the standard deviation of inflation.

line is instead the correlation measured in the data, with a 90 percent bootstrap interval around this estimate represented by the dashed lines. The serial correlation of output growth in the model is very close to its empirical counterpart and well within the data-uncertainty band. For inflation and the interest rate, the model serial correlations are on the high end of the band. This excessive persistence is a result of the low-frequency component in both variables associated with the inflation target, as we can infer from the variance decomposition in Table 2.

According to the model, shocks to the inflation target account for 85 percent of the unconditional variance of inflation and 38 percent of that of the nominal interest rate. Although we do not calculate a variance decomposition by frequency, we know that the contribution of the inflation target shock is concentrated at low frequencies, since this shock is very persistent (the posterior mean of its autocorrelation coefficient is 0.98). This finding suggests that the model faces a trade-off between accommodating the downward drift in inflation in the first part of our sample and providing a more balanced account of the sources of inflation variability.

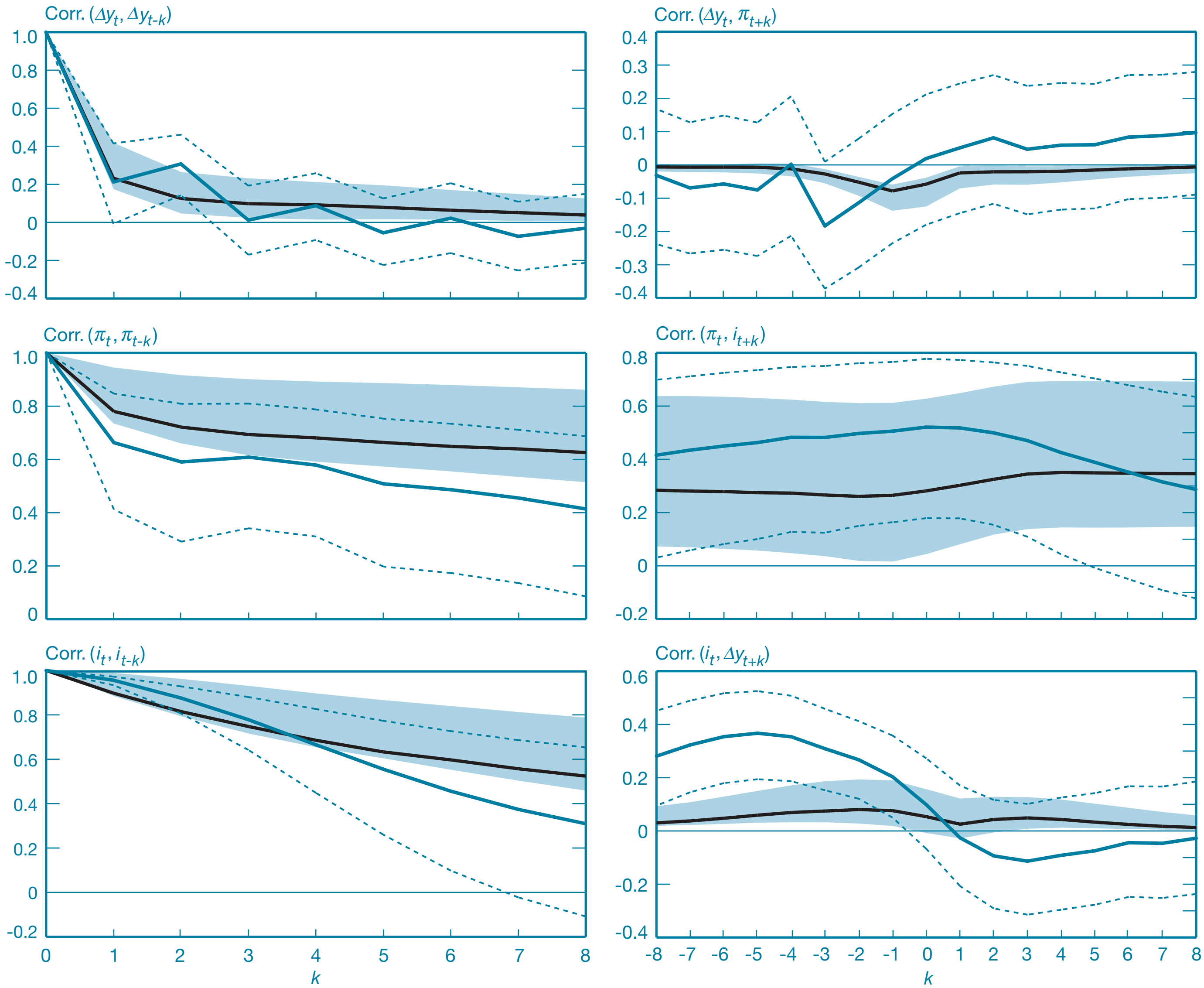
The rest of the variance decomposition accords well with conventional wisdom. The productivity shock plays an

According to the model, shocks to the inflation target account for 85 percent of the unconditional variance of inflation and 38 percent of that of the nominal interest rate. This finding suggests that the model faces a trade-off between accommodating the downward drift in inflation in the first part of our sample and providing a more balanced account of the sources of inflation variability.

important role in accounting for the volatility of output growth, although the demand shock and the monetary policy shocks (interest rate plus inflation target) are also non-negligible. Moreover, the demand shock accounts for more than half of the variance of the nominal interest rate. Finally, mark-up shocks play a minor role as sources of volatility. This finding has potentially important policy implications, since in the model mark-up shocks are the only source of the aforementioned policy trade-off between inflation and real activity.

Returning now to the cross-correlations in Chart 2, we see that the model is quite successful in capturing the lead-lag relationships in the data. In our sample, there is no statistically significant predictability of future inflation through current output *growth*, a pattern that is reproduced by the model. The

Chart 2
Correlations



Source: Authors' calculations.

Notes: The black line represents the median model-implied correlation across draws from the posterior; the shaded area represents the interval between the 5th and 95th percentiles across those same draws. The solid blue line represents the median autocorrelation across bootstrap replications of a VAR(4) fit to the data; the dashed blue lines represent the interval between the 5th and 95th percentiles across those same replications. Each statistic is calculated at horizons $k = 0, \ldots 8$ for autocorrelations and at horizons $k = -8, \ldots, 0, \ldots 8$ for cross-correlations.

model also reproduces the positive correlation between inflation and the nominal interest rate present in the data both in the leads and in the lags (the middle right panel of the chart). The positive correlation between current interest rates and future inflation might seem puzzling at first. We would expect higher interest rates to bring inflation down over time, which should make the correlation negative. However, over our sample, this negative relationship is confounded by the low-frequency positive comovement between inflation and the nominal interest rate induced by the Fisher effect. Recall that inflation and the nominal interest rate in fact are persistently above their unconditional sample average over the first third of the sample and are persistently below it after about 1992.

The bottom right panel of Chart 2 reports the dynamic

Table 2
Variance Decomposition
Percent

	Shocks				
Variable	Demand	Productivity	Mark-Up	Interest Rate	Inflation Target
GDP growth	0.20	0.56	0.07	0.04	0.13
	[.13, .28]	[.45, .67]	[.04, .10]	[.02, .06]	[.07, .19]
Core PCE inflation	0.04	0.01	0.10	0.00	0.85
	[.01, .06]	[.00, .02]	[.04, .17]	[.00, .01]	[.76, .94]
Federal funds rate	0.54	0.06	0.01	0.01	0.38
	[.32, .76]	[.00, .13]	[.00, .02]	[.01, .02]	[.16, .61]

Source: Authors' calculations.

Notes: The table reports the share of the unconditional variance of each observable variable contributed by each shock. The point estimates are medians across draws from the posterior; the 5th and 95th posterior percentiles across those same draws are in brackets. PCE is personal consumption expenditures.

correlation between output growth and the nominal interest rate. In the data, high growth rates of output predict high nominal interest rates one to two years ahead, but this predictability is much less pronounced in the model. Moreover, this discrepancy is statistically significant in the sense that the model-implied median autocorrelation lies outside the 90 percent bootstrap interval computed from the data. This failure to match the data highlights the main empirical weakness of our model: its demand-side specification. As in most of the DSGE literature, our demand block consists of the Euler equation of a representative consumer. Standard specifications of a Euler equation of the type adopted here provide an inaccurate description of the observed relationship between the growth rate of consumption (or output, as in our case) and financial returns, including interest rates, as first documented by Hansen and Singleton (1982, 1983) and subsequently confirmed by many others (see Campbell [2003] for a review). Improving the performance of the current generation of DSGE models in this dimension would be an important priority for future research.

We now report our estimates of a few of the variables that play an important role in the model, but that are not directly observable. We focus on the three latent variables that enter the interest rate rule: the inflation target π_t^*, the output gap $y_t - y_t^e$, and the efficient real interest rate r_t^e (Chart 3). As in Charts 1 and 2, the black line is the median estimate across draws from the model's posterior and the shaded area represents a 90 percent posterior probability interval.

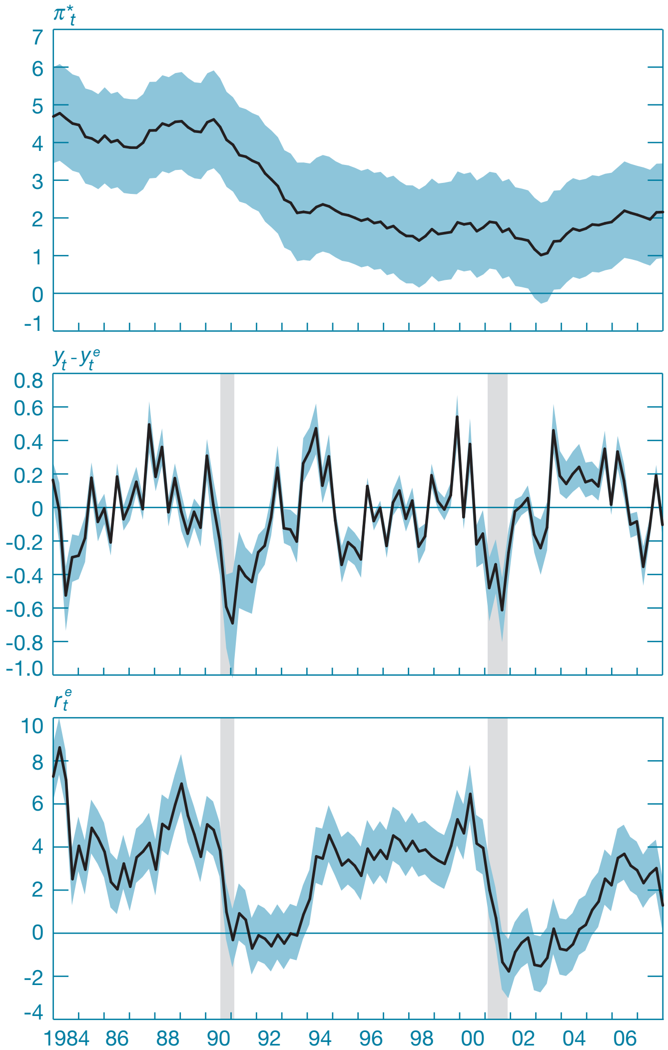
Starting from the top panel, we note that the estimated inflation target captures well the step-down in inflation from a local mean above 4 percent between 1984 and 1991 to an average value of around 2 percent since 1994. This permanent reduction in inflation represents the last stage of the disinflation process initiated by Federal Reserve Chairman Volcker in 1979, which became known as an example of opportunistic disinflation (Orphanides and Wilcox 2002). Needless to say, the estimated target is not completely smooth, but it also displays some higher frequency variation. For example, it reaches a minimum of around 1 percent at the beginning of 2003, but moves closer to 2 percent over 2004. (The next section studies in more detail the implications of these movements.)

The middle and bottom panels of Chart 3 report estimates of the output gap—the percentage deviation of output from its efficient level—and of the efficient real interest rate. Several features of the estimated output gap are noteworthy. First, its two deepest negative troughs correspond to the two recessions in our sample. In this respect, our model-based output gap conforms well with more conventional measures of this variable, such as the one produced by the Congressional Budget Office (CBO). However, the shortfall of output from its efficient level is never larger than 0.7 percent, even in these recessionary episodes. By comparison, the CBO output gap is as low as -3.2 percent in 1991. The amplitude of the business cycle fluctuations in our estimated output gap is small because the efficient level of output is a function of all the shocks in the model and therefore it tracks actual output quite closely. The last notable feature of the efficient output gap is that it displays a very pronounced volatility at frequencies higher than the business cycle. During the 1990s expansion, for example, it crosses the zero line about a dozen times.

Compared with the output gap, the efficient real rate is significantly smoother. Although some high-frequency variation remains, the behavior of the efficient real rate is dominated by swings at the business cycle frequency. The rate spikes and then plunges for some time before the onset of

Chart 3

Kalman Smoother Estimates of Latent Variables



Source: Authors' calculations.

Notes: The black line represents the Kalman smoother estimate of the relevant latent variable conditional on the posterior mean of the parameters; the shaded area represents the interval between the 5th and 95th percentiles of the Kalman smoother estimates across draws from the posterior. The vertical bands indicate NBER recessions.

recessions and recovers a few quarters into the expansions. It is interesting to note that the efficient real rate was negative for the entire period between 2001 and 2004—a time when the FOMC was concerned about the possibility that the U.S. economy would fall into a liquidity trap.[12] A negative efficient real interest rate is a necessary condition for the zero bound on nominal interest rates to become binding, and hence for the liquidity trap to become a problem.

[12] A liquidity trap describes a situation in which nominal interest rates have reached their zero lower bound, as in Japan in the 1990s, and therefore cannot be lowered any further.

5. The Model at Work: The Pickup in Inflation in the First Half of 2004

To show how our model can be used to address specific policy questions, we examine a particular historical episode: the puzzling pickup in inflation in the first half of 2004. This exercise allows us to illustrate how we use the model's forecasts to construct alternative scenarios for counterfactual policy analysis. Moreover, our analysis offers potentially interesting lessons for the current situation—although inflation has recently been quite low, there has been some concern that it might accelerate in the near future.

After approaching levels close to 1 percent between 2002 and 2003, core PCE inflation started moving higher in mid-2003. This pickup accelerated significantly in the first half of 2004, when (year-over-year) core inflation moved from about 1.5 percent to more than 2 percent, where it remained until the end of 2008. We use our DSGE model to analyze the sources of this unusually rapid and persistent step-up in the level of inflation.

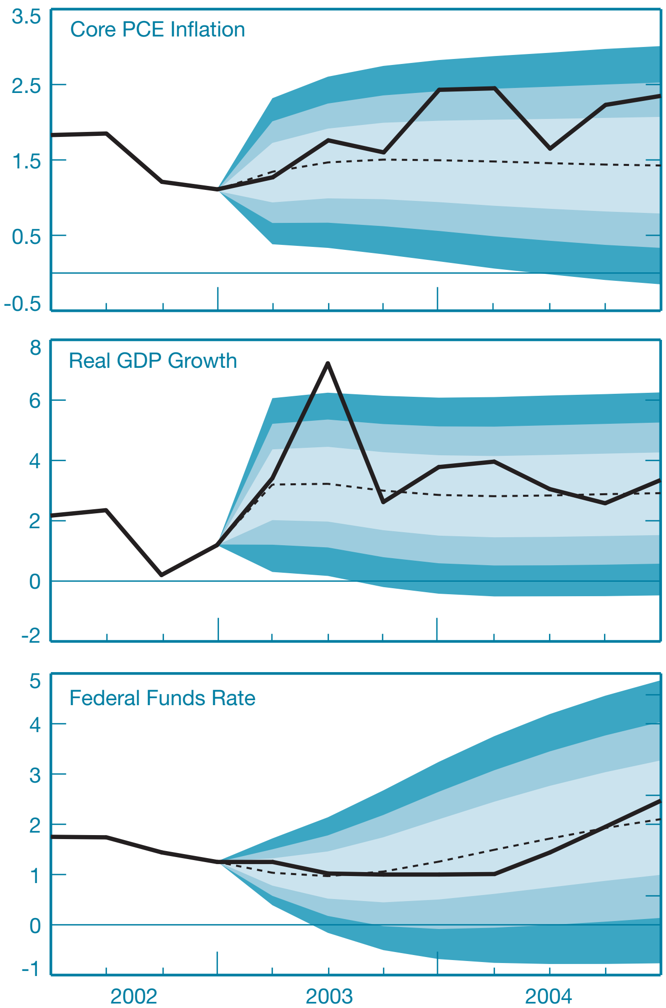
To show how our model can be used to address specific policy questions, we examine a particular historical episode: the puzzling pickup in inflation in the first half of 2004. This exercise allows us to illustrate how we use the model's forecasts to construct alternative scenarios for counterfactual policy analysis.

We organize our discussion around three questions. First, was the surge in inflation forecastable? As we will see, the answer to this question is no, at least from the perspective of our model. Second, what accounts for the discrepancy between the model's forecast and the observed paths of inflation, output growth, and the federal funds rate? Third, could monetary policy have achieved a smooth transition to inflation rates below 2 percent and, if so, at what cost in terms of added volatility in output and the interest rate?

Chart 4 presents forecasts of quarterly core PCE inflation, real GDP growth, and the federal funds rate from the DSGE model. The forecast starts in 2003:1, when quarterly inflation reached 1.1 percent (at an annual rate)—its lowest level following the 2001 recession—and extends through the beginning of 2005. In each panel, the dashed line represents

Chart 4

Forecasts of Observable Variables



Source: Authors' calculations.

Notes: The dashed line represents the forecast of the relevant variable conditional on the posterior mean of the parameters; the solid line represents the observed realization. The shaded areas represent 50 (light blue), 75 (medium blue), and 90 percent (dark blue) symmetric probability intervals for the forecast at each horizon. PCE is personal consumption expenditure.

the expected value of the forecast, while the bands show the 50 (light blue), 75 (medium blue), and 90 (dark blue) percent probability intervals. The solid line shows the realized data.

The model performs well in its forecast of output and the federal funds rate, especially in the medium term. Inflation, by comparison, is close to the mean forecast in 2003, but is well above it in 2004 and beyond. Moreover, the probability intervals for the forecast suggest that this realization of inflation was quite unusual, as we see from the fact that the solid line borders the 75 percent probability interval in the first half of 2004. This means that in 2003:1, the model would have assessed only about one in ten chances (12.5 percent) of inflation being as high as it was in that period.

From an economic perspective, it is interesting to note that these sizable forecast errors for inflation roughly correspond to the "considerable period" era that extended from June 2003 to June 2004. At that time, the FOMC kept the federal funds rate constant at 1 percent to guard against the risk of deflation, while indicating in its statement that "policy accommodation can be maintained for a considerable period."[13] According to the model's projection, this path for the federal funds rate represents a deviation from the policy stance historically maintained by the Federal Reserve in similar macroeconomic circumstances. Based on the estimated interest rate rule, in fact,

The DSGE forecast is just a description of what would happen to the variables of interest if we allowed the model economy to "run" from its initial condition, without introducing any innovations. Any observed deviation from the forecast, therefore, must be attributable to the realization of a particular combination of such innovations.

the DSGE predicts a slow rise in the interest rate over 2003 and 2004. Instead, the FOMC maintained the federal funds rate at 1 percent through the first half of 2004.

However, the pickup in inflation over this period is significantly more "unusual" than the deviation of the federal funds rate from the historical norm. Actual inflation in 2004 is mostly outside the 50 percent probability interval of the model forecast (the light blue band), while the actual federal funds rate remains well within it. Moreover, the acceleration in inflation is not accompanied by unexpectedly high real growth, suggesting that it cannot be fully explained by the traditional channel of transmission from an overheated economy to higher inflation.

What else, then, accounts for the unexpected and unlikely deviation of inflation from the model's forecast over 2004? The DSGE framework provides a particularly useful way of addressing this question. As we discuss in Section 2, the economic outcomes predicted by the model—the levels of inflation, output, and the interest rate—are the result of the endogenous responses of the agents in the economy to the

[13] This formulation was maintained in the FOMC statement from August 2003 to December 2003, and was later substituted with "policy accommodation can be removed at a pace that is likely to be measured."

Chart 5
Forecasts of Shocks

Source: Authors' calculations.

Notes: The dashed line represents the forecast of the relevant shock conditional on the posterior mean of the parameters while the solid line represents an estimate of the realization based on the Kalman smoother. The shaded area represents the 75 percent symmetric probability interval for the forecast at each horizon.

realization of a set of exogenous processes, such as productivity or desired mark-ups. The innovations to these driving processes account for the deviations of the data from the model's forecast. In fact, the DSGE forecast is just a description of what would happen to the variables of interest if we allowed the model economy to "run" from its initial condition, without introducing any innovations. Any observed deviation from the forecast, therefore, must be attributable to the realization of a particular combination of such innovations.[14]

Chart 5 depicts the combinations of exogenous driving processes that, according to the estimated DSGE model, are responsible for the observed path of inflation, output, and the interest rate over the period we analyze. In each panel, the dashed line represents the evolution of the shock associated with the mean forecast while the solid line represents the sequence of shocks corresponding to the actual realization of the observable variables. As in Chart 4, the medium blue band denotes the 75 percent probability interval for the forecast.

[14] In this study, we distinguish between exogenous driving processes—shocks, for short—and innovations. Driving processes can be autocorrelated, and thus forecastable, while their innovations are i.i.d.

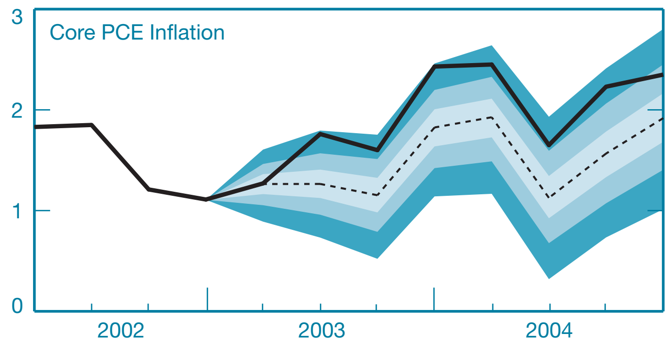
The contribution of three shocks stands out. First, the demand shock recovers from almost -4 percent to around -1 percent. This movement is particularly pronounced during 2004, when inflation was picking up. However, this profile is broadly consistent with the shock's expected evolution, represented by the dashed line. The productivity shock is also broadly in line with expectations, with the exception of 2003:3; this spike in productivity accounts for the corresponding spike in output growth in that quarter.

However, the most significant and direct contribution to the surge in inflation comes from a sizable upward movement in the inflation target, π_t^*. According to our estimates, this target moves by about 1 percentage point, from less than 1 percent to close to 2 percent. Moreover, this movement is at the edge of the 75 percent probability interval for the forecast, suggesting that the realization of this driving process is indeed quite unusual.

To quantify more directly the effect on inflation of the unexpected increase in the implicit inflation target, we depict what would have happened to core PCE inflation in the absence of such an increase (Chart 6). Here, the solid line

Chart 6

Conditional Forecast of Inflation



Source: Authors' calculations.

Notes: The dashed line represents a forecast of inflation conditional on the Kalman smoother estimates of all shocks except for those to the inflation target; the solid line represents the observed realization. The shaded areas represent 50 (light blue), 75 (medium blue), and 90 percent (dark blue) symmetric probability intervals for this conditional forecast. Therefore, they represent uncertainty stemming from future realizations of the inflation target shock alone. PCE is personal consumption expenditure.

is realized inflation. The dashed line represents the counterfactual path of inflation predicted by the model in the absence of shocks to the inflation target. In other words, this is a forecast for inflation, conditional on the estimated path of all but the inflation target shock. The bands therefore represent the usual probability intervals, but in this case they are computed around this conditional forecast.

The chart confirms our conclusion on the role of innovations to the inflation target in accounting for the observed evolution of inflation. According to the model, core inflation would not have increased to above 2 percent, as it did for most of 2004, without the steady increase in the inflation target over the same period. In fact, inflation would have remained within the "comfort zone" of 1 to 2 percent. Moreover, note that the solid line of realized inflation is mostly inside the area associated with the 90 percent probability interval for the conditional forecast. This suggests that the share of the forecast error in inflation accounted for by the innovations in the inflation target in this episode is unusually large compared with the historical average. This is just a more formal way of saying that the increase in the inflation target is disproportionately responsible for the observed increase in inflation that we examine.

The estimated rise in the implicit inflation target provides the missing link for a unified explanation of the pickup in inflation, the "considerable period" monetary policy, and the absence of a concomitant acceleration in output growth. In the model, the inflation target is the main driver of movements in inflation expectations, which are a key determinant of firms' pricing behavior together with the amount of slack in the economy. According to the DSGE model, therefore, a significant fraction of the inflation acceleration in 2003-04 can be attributed to a change in inflation expectations, driven by an increase in the Federal Reserve's implicit inflation target as perceived by the private sector. This increase in the perceived target in turn is consistent with the unusually loose stance of monetary policy maintained by the FOMC during the "considerable period" era.

This brings us to the third question: If the DSGE model is correct, and the pickup in inflation in 2004 is attributable to an increase in the implicit inflation target perceived by the public, could the Federal Reserve have prevented this development?

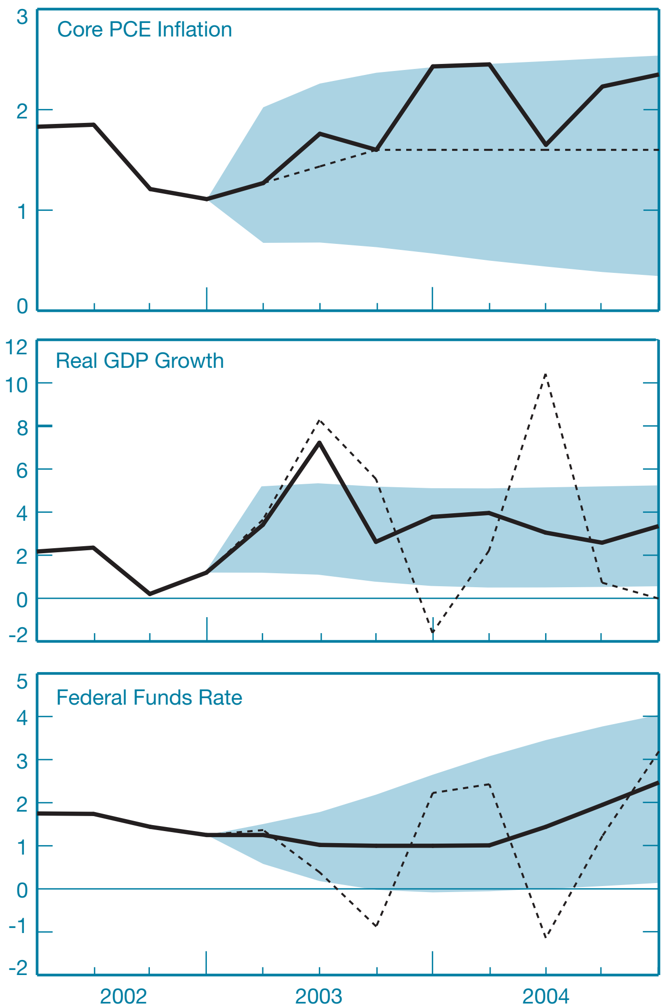
Charts 7 and 8 show the results of this counterfactual analysis. Both charts display the data (solid line) along with the counterfactual outcomes for the economy predicted by the model under a policy consistent with the stabilization of core inflation at 1.6 percent through 2004. The way in which this policy is implemented, however, is different in the two cases. In Chart 7, we present the outcomes associated with what we call a "no-communication" monetary strategy (dashed line) while in Chart 8 we compare these results with those that would emerge under a "full-communication" strategy (blue line). Under the no-communication strategy, the path for the interest rate compatible with the desired evolution of inflation is achieved each period through "surprise" departures from the historical rule. In contrast, under the full-communication strategy, the Federal Reserve implements the same path for inflation by announcing an inflation target that is consistent with it.[15]

According to the DSGE model . . . a significant fraction of the inflation acceleration in 2003-04 can be attributed to a change in inflation expectations, driven by an increase in the Federal Reserve's implicit inflation target as perceived by the private sector.

[15] Technically, in both cases we choose shocks to the monetary policy rule that are compatible with the desired evolution of inflation, conditional on the smoothed value of all other disturbances. Under the no-communication strategy, the shocks we choose are the i.i.d. monetary shocks, ε_t^i. Under the full-communication strategy, our chosen shocks are to the inflation target, $\varepsilon_t^{\pi^*}$.

Chart 7

"No-Communication" Counterfactual

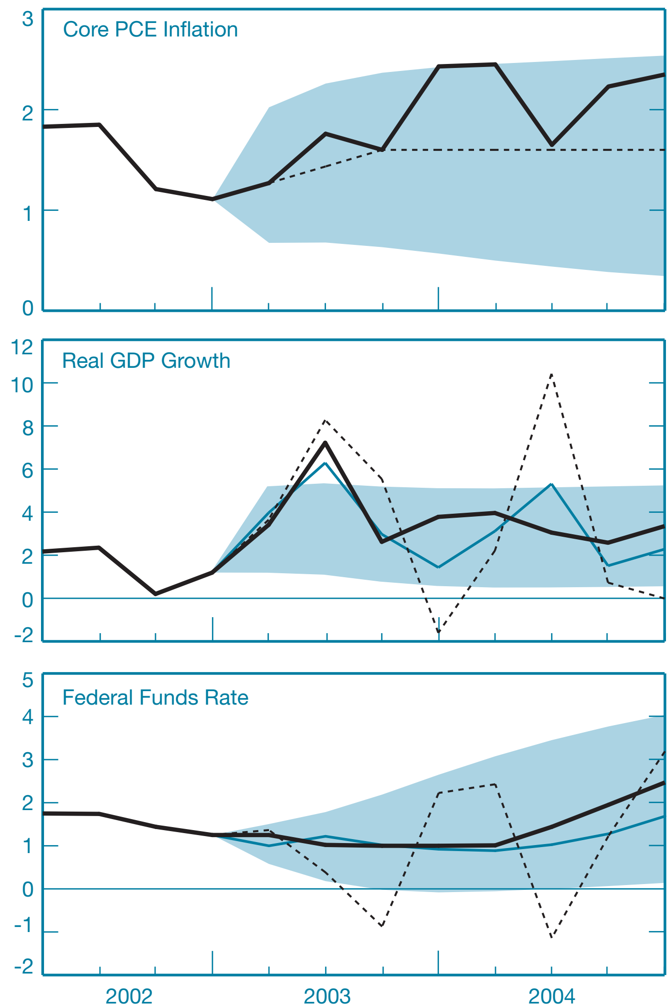


Source: Authors' calculations.

Notes: The dashed line represents the counterfactual evolution of the economy predicted by our model had monetary policy been set to achieve the path for inflation depicted in the top panel. This counterfactual is conditional on the posterior mean of the parameters. Under the "no-communication" scenario, the desired path for inflation is achieved by the choice of the interest rate shock while all other shocks are set at their Kalman smoother estimate. The shaded area represents the 75 percent symmetric probability interval for the unconditional forecast, which is the same as in Chart 4. The black line represents the observed realization of each series. PCE is personal consumption expenditure.

Chart 8

"Full-Communication" Counterfactual



Source: Authors' calculations.

Notes: The blue line represents the counterfactual evolution of the economy predicted by our model had monetary policy been set to achieve the path for inflation depicted in the top panel. This counterfactual is conditional on the posterior mean of the parameters. Under the "full-communication" scenario, the desired path for inflation is achieved by the choice of the inflation target while all other shocks are set at their Kalman smoother estimate. The shaded area represents the 75 percent symmetric probability interval for the unconditional forecast, which is the same as in Chart 4. The black line represents the observed realization of each series. The dashed line is the conditional forecast under the "no-communication" scenario. PCE is personal consumption expenditure.

The crucial difference between the results obtained under the two scenarios stems from the key role that expectations play in the DSGE model. Under the full-communication strategy, inflation expectations are immediately affected by the announcement of an inflation target. These expectations in turn have a direct effect on actual inflation without requiring a contraction in real activity to force businesses to contain their price increases. Under the no-communication strategy, inflation expectations remain at their historical level. As a result, inflation can be controlled only by increasing interest rates to contain GDP growth.

The way in which we model the full-communication scenario is quite stark. In practice, expectations would be unlikely to adjust instantaneously, even if the Federal Reserve were completely transparent about its inflation target. Nevertheless, the differences between the results of the two policy strategies are striking. In the no-communication case, inflation can be stabilized only through wild movements in the federal funds rate. As a result, GDP growth is also extremely volatile: it falls below zero in 2004:1, but then recovers to a quarterly (annualized) growth rate of 10 percent and ends the period at zero. These movements in output are indeed extreme. They lie well outside the 75 percent forecast probability interval reported in the chart. In fact, the quantitative details of the evolution of output and the interest rate under the counterfactual simulations should not be taken literally, since they depend significantly on the details of the model and on the assumption that the central bank insists on perfectly stabilizing current inflation. However, the qualitative pattern of higher volatility under the no-communication strategy is a robust feature of models in which expectations matter.

Our analysis might be interpreted as supportive of the policy stance adopted by the central bank in 2003-04 as part of a successful preemptive strike against a liquidity trap.

Under the full-communication strategy, in contrast, the desired path for inflation can be achieved with much less pronounced fluctuations in real growth and an almost unchanged policy relative to the actual path. Interest rates need not rise and output need not fall significantly because a shift in expectations brought about by clear communication of the Federal Reserve's inflation objective largely brings inflation under control.

Note that the results of these counterfactual exercises should be interpreted with caution. Their objective is not to prescribe an alternative to the policy followed in 2004, but rather to investigate how a different path for inflation might have been achieved. In fact, according to Krugman (1998) and Eggertsson and Woodford (2003), an increase in inflation expectations might be the best monetary strategy to escape a liquidity trap. Many have argued that the main objective of the Federal Reserve around 2003 was to minimize the U.S. economy's likelihood of falling into such a trap.[16] From this perspective, our analysis might be interpreted as supportive of the policy stance adopted by the central bank in 2003-04 as part of a successful preemptive strike against a liquidity trap.

6. Conclusion

This article provides an introduction to dynamic stochastic general equilibrium models and presents an example of their use as tools for monetary policy analysis. Given the mainly educational nature of our presentation, we simplify by using a small-scale model designed to account for the behavior of three key macroeconomic variables: GDP growth, core PCE inflation, and the federal funds rate. Despite its simplicity, our model is rich enough to reproduce some of the salient features of the series of interest. It also allows us to highlight the components common to more articulated and realistic DSGE specifications.

Our model offers insight into the causes of the abrupt pickup in inflation in the first half of 2004, from levels close to 1 percent at the beginning of 2003 to values steadily above 2 percent through the end of 2008. This exercise highlights the central role of expectations in the transmission of shocks and policy impulses in DSGE models. The main lesson that we derive from the exercise is that the most effective approach to controlling inflation is through the management of expectations, rather than through actual movements of the policy instrument. This lesson seems to be well understood by the public, given the amount of attention and speculation that usually surround the pronouncements of central bankers on their likely future actions. DSGE models have the potential to broaden this understanding by adding a quantitative assessment of the link between current policy, expectations, and economic outcomes—and thus to clarify the effect that different systematic approaches to policy have on those outcomes.

[16] In its August 2003 statement, the FOMC observed that "on balance, the risk of inflation becoming undesirably low is likely to be the predominant concern for the foreseeable future." Very low or negative levels of inflation are one of the most likely triggers of a liquidity trap.

Technical Appendix

The table reports information on the prior distribution for the parameters of the model. Further details on the parameters and the structure of the model are available from the corresponding authors.

Parameter	Distribution	Mean	Standard Deviation
β	Calibrated	0.99	—
ξ	Gamma	0.1	0.05
ω	Gamma	1.0	0.2
η	Beta	0.6	0.2
ζ	Beta	0.6	0.2
ρ	Beta	0.7	0.15
ϕ_π	Normal	1.5	0.25
ϕ_y	Normal	0.5	0.2
π^*	Normal	2.0	1.0
r	Normal	2.0	1.0
γ	Normal	3.0	0.35
ρ_{π^*}	Beta	0.95	0.04
ρ_δ	Beta	0.5	0.2
ρ_γ	Beta	0.5	0.2
ρ_u	Beta	0.5	0.2
σ_δ	InvGamma	0.5	2.0
σ_γ	InvGamma	0.5	2.0
σ_u	InvGamma	0.5	2.0
σ_{π^*}	InvGamma	0.2	1.0
σ_i	InvGamma	0.5	2.0

References

Adolfson, M., M. K. Andersson, J. Lindé, M. Villani, and A. Vredin. 2007. "Modern Forecasting Models in Action: Improving Macroeconomic Analyses at Central Banks." International Journal of Central Banking 3, no. 4 (December): 111-44.

Adolfson, M., S. Laséen, J. Lindé, and M. Villani. 2008. "Evaluating an Estimated New Keynesian Small Open-Economy Model." Journal of Economic Dynamics and Control 32, no. 8 (August): 2690-2721.

An, S., and F. Schorfheide. 2007. "Bayesian Analysis of DSGE Models." Econometric Reviews 26, nos. 2-4 (April): 113-72.

Bernanke, B. S., M. Gertler, and S. Gilchrist. 1999. "The Financial Accelerator in a Quantitative Business Cycle Framework." In J. B. Taylor and M. Woodford, eds., Handbook of Macroeconomics, vol. 1, 1341-93. Elsevier.

Bils, M., and P. J. Klenow. 2004. "Some Evidence on the Importance of Sticky Prices." Journal of Political Economy 112, no. 5 (October): 974-85.

Calvo, G. A. 1983. "Staggered Prices in a Utility-Maximizing Framework." Journal of Monetary Economics 12, no. 3 (September): 383-98.

Campbell, J. Y. 2003. "Consumption-Based Asset Pricing." In G. M. Constantinides, M. Harris, and R. M. Stulz, eds., Handbook of the Economics of Finance, vol. 1, 803-87. North-Holland.

Campbell, J. Y., and J. H. Cochrane. 2000. "Explaining the Poor Performance of Consumption-Based Asset Pricing Models." Journal of Finance 55, no. 6 (December): 2863-78.

Christiano, L., R. Motto, and M. Rostagno. 2007. "Financial Factors in Business Cycles." Unpublished paper.

Clarida, R. G., J. Galí, and M. Gertler. 1999. "The Science of Monetary Policy: A New Keynesian Perspective." Journal of Economic Literature 37, no. 4 (December): 1661-1707.

Curdia, V., A. Ferrero, and A. Tambalotti. 2009. "Which Output Gap? Evaluating Interest Rate Rules in an Estimated DSGE Model." Unpublished paper, Federal Reserve Bank of New York.

Curdia, V., and M. Woodford. 2009. "Conventional and Unconventional Monetary Policy." Federal Reserve Bank of New York Staff Reports, no. 404, November.

Dixit, A. K., and J. E. Stiglitz. 1977. "Monopolistic Competition and Optimum Product Diversity." American Economic Review 67, no. 3 (June): 297-308.

Edge, R. M., M. T. Kiley, and J.-P. Laforte. 2008. "Natural Rate Measures in an Estimated DSGE Model of the U.S. Economy." Journal of Economic Dynamics and Control 32, no. 8 (August): 2512-35.

Eggertsson, G. B., and M. Woodford. 2003. "The Zero Bound on Interest Rates and Optimal Monetary Policy." Brookings Papers on Economic Activity, no. 1: 139-211.

Eichenbaum, M. S., L. Hansen, and K. J. Singleton. 1988. "A Time Series Analysis of Representative Agent Models of Consumption and Leisure Choice under Uncertainty." Quarterly Journal of Economics 103, no. 1 (February): 51-78.

Erceg, C. J., L. Guerrieri, and C. Gust. 2006. "SIGMA: A New Open-Economy Model for Policy Analysis." International Journal of Central Banking 2, no. 1 (March): 1-50.

Galí, J., and M. Gertler. 1999. "Inflation Dynamics: A Structural Econometric Analysis." Journal of Monetary Economics 44, no. 2 (October): 195-222.

———. 2007. "Macroeconomic Modeling for Monetary Policy Evaluation." Journal of Economic Perspectives 21, no. 4 (fall): 25-45.

Gerali, A., S. Neri, L. Sessa, and F. Signoretti. 2008. "Credit and Banking in a DSGE Model." Unpublished paper, Bank of Italy, June.

Gertler, M., and P. Karadi. 2009. "A Model of Unconventional Monetary Policy." Unpublished paper.

Gertler, M., L. Sala, and A. Trigari. 2008. "An Estimated Monetary DSGE Model with Unemployment and Staggered Nominal Wage Bargaining." Journal of Money, Credit, and Banking 40, no. 8 (December): 1713-64.

References (Continued)

Gilchrist, S., A. Ortiz, and E. Zakrajsek. 2009. "Credit Risk and the Macroeconomy: Evidence from an Estimated DSGE Model." Unpublished paper.

Hansen, L. P., and K. J. Singleton. 1982. "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models." Econometrica 50, no. 5 (September): 1269-86.

———. 1983. "Stochastic Consumption, Risk Aversion, and the Temporal Behavior of Asset Returns." Journal of Political Economy 91, no. 2 (April): 249-65.

Krugman, P. 1998. "It's Baaack! Japan's Slump and the Return of the Liquidity Trap." Brookings Papers on Economic Activity, no. 2: 137-87.

Nakamura, E., and J. Steinsson. 2008. "Five Facts about Prices: A Reevaluation of Menu Cost Models." Quarterly Journal of Economics 123, no. 4 (November): 1415-64.

Orphanides, A., and D. W. Wilcox. 2002. "The Opportunistic Approach to Disinflation." International Finance 5, no. 1 (spring): 47-71.

Sbordone, A. M. 2002. "Prices and Unit Labor Costs: A New Test of Price Stickiness." Journal of Monetary Economics 49, no. 2 (March): 265-92.

Smets, F., and R. Wouters. 2003. "An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area." Journal of the European Economic Association 1, no. 5 (September): 1123-75.

———. 2007. "Shocks and Frictions in U.S. Business Cycles: A Bayesian DSGE Approach." American Economic Review 97, no. 3 (June): 586-606.

Taylor, J. B. 1993. "Discretion versus Policy Rules in Practice." Carnegie-Rochester Conference Series on Public Policy 39, December: 195-214.

———. 2008. "The Costs and Benefits of Deviating from the Systematic Component of Monetary Policy." Remarks delivered at the Federal Reserve Bank of San Francisco Conference on Monetary Policy and Asset Markets, San Francisco, California, February 22.

Woodford, M. 2003. Interest and Prices: Foundations of a Theory of Monetary Policy. Princeton, N.J.: Princeton University Press.

Kenneth D. Garbade, Frank M. Keane, Lorie Logan, Amanda Stokes, and Jennifer Wolgemuth

The Introduction of the TMPG Fails Charge for U.S. Treasury Securities

- Prior to May 2009, market convention enabled a seller of Treasury securities to postpone—without any explicit penalty and at an unchanged invoice price—its obligation to deliver the securities.
- The September 2008 insolvency of Lehman Brothers exposed a flaw in this convention, when a decline in short-term interest rates set the stage for an extraordinary volume of settlement fails that threatened to erode the perception of the market as being free of credit risk.
- In response, the Treasury Market Practices Group introduced a "dynamic fails charge" for Treasury securities in May 2009.
- The fails charge incentivizes timely settlement by providing that a buyer of Treasury securities can claim monetary compensation from a seller if the seller fails to deliver on a timely basis.
- The fails charge mitigated a key dysfunctionality in the market and illustrates the value of public and private sector cooperation in responding to altered market conditions.

1. Introduction

Securities transactions commonly involve a variety of market conventions—widely accepted ways of doing business that persist through time even though not mandated by law or regulation. Commonplace examples include the quotation of prices for Treasury bonds in increments of 32nds (and fractions of a 32nd) of a percent of principal value (rather than in decimal increments) and the quotation of Treasury bills in terms of discount rates (rather than prices or yields).

In most cases, market conventions are useful or, at worst, innocuous. In some cases, however, a new use for an old instrument can render a convention in need of revision. One particularly notorious example was the convention—observed prior to 1982—of ignoring accrued interest on Treasury bonds sold on repurchase agreements (also known as repos, or RPs). The convention made sense as long as repos were used primarily to borrow money from creditworthy lenders that held the bonds simply to limit their exposure to credit risk. It made less sense when highly leveraged securities dealers began to use repos to borrow bonds to deliver on short sales. The 1982 Drysdale episode illuminated the risks involved in ignoring accrued interest and prompted the Federal Reserve Bank of New York to orchestrate a change in the convention.[1]

[1] Garbade (2006).

Kenneth D. Garbade is a senior vice president, Frank M. Keane an assistant vice president, Lorie Logan a vice president, Amanda Stokes a trader/analyst, and Jennifer Wolgemuth a counsel and assistant vice president at the Federal Reserve Bank of New York.
Correspondence: kenneth.garbade@ny.frb.org

The authors are grateful to Murray Pozmanter and Cartier Stennis for assistance in researching this article and to two anonymous referees for helpful suggestions. They are especially grateful to the members of the Treasury Market Practices Group, identified in Box 4, for the time and effort they devoted to making the fails charge a reality, as well as to all of the other individuals, identified in the appendix, who provided invaluable feedback and assistance during the implementation of the fails charge. 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.

A market convention may also require revision following a change in the economic environment. This article discusses a recent example: the convention of postponing—without any explicit penalty and at an unchanged invoice price—a seller's obligation to deliver Treasury securities if the seller fails to deliver the securities on a scheduled settlement date. As discussed in more detail below, as long as short-term interest rates were above about 3 percent, the time value of money usually sufficed to incentivize timely settlement of transactions in Treasury securities. However, when short-term rates fell to near zero following the insolvency of Lehman Brothers Holdings Inc. in September 2008, the time value of money no longer provided adequate incentive and the Treasury market experienced an extraordinary volume of settlement fails. Both the breadth of the fails across a large number of securities and the persistence of the fails were unprecedented and threatened to erode the perception of the Treasury market as a market free of credit risk.[2] In response, the Treasury Market Practices Group (TMPG)—a group of market professionals committed to supporting the integrity and efficiency of the U.S. Treasury market—worked over a period of six months to revise the market convention for settlement fails, developing a "dynamic fails charge" that, when short-term interest rates are *below* 3 percent, produces an economic incentive to settle trades roughly equivalent to the incentive that exists when rates are *at* 3 percent. Thus, the TMPG fails charge preserves a significant economic incentive for timely settlement even when interest rates are close to zero.

A market convention may . . . require revision following a change in the economic environment. This article discusses a recent example: the convention of postponing—without any explicit penalty and at an unchanged invoice price—a seller's obligation to deliver Treasury securities if the seller fails to deliver the securities on a scheduled settlement date.

[2] See, for example, Wrightson, *Federal Reserve Data*, October 17, 2008 ("The breakdown in the clearing mechanism for the Treasury market is beginning to emerge as a top-tier policy concern. The safe-haven status of Treasury securities is one of the few advantages the government market has left in a year in which net Treasury borrowing needs . . . are likely to exceed $1 trillion by a large margin. At some point, though, buyers will think twice about buying a 'safe-haven' asset for peace of mind if they have doubts about their counterparty's ability to deliver the security.").

This article describes the introduction of the TMPG fails charge. The introduction of the fails charge is important for two reasons. First, it mitigated an important dysfunctionality in a market of critical significance both to the Federal Reserve in its execution of monetary policy and to the country as a whole. Second, it exemplified the value of cooperation between the public and private sectors in responding to altered market conditions in a flexible, timely, and innovative fashion.

Our study is divided into three parts. The first part (Sections 2-5) describes settlement processes and settlement fails in the Treasury market, explains why sellers usually try to avoid fails, describes industry and Federal Reserve efforts to mitigate settlement fails prior to 2008, and briefly reviews three episodes of chronic fails in the Treasury market. The second part (Section 6) describes the tsunami of fails that followed Lehman's insolvency. The balance of the study (Sections 7-10) explains the TMPG's response. Section 11 concludes.

2. Settlements and Settlement Fails in U.S. Treasury Securities

A transaction in Treasury securities is said to "settle" when the seller delivers the securities to, and receives payment from, the buyer. The two most important settlement processes are bilateral settlement and multilateral net settlement. Before describing those processes, we explain how market participants establish and transfer ownership of Treasury securities.

2.1 Establishing and Transferring Ownership of Treasury Securities

For more than three decades, investors have established ownership of Treasury securities through Federal Reserve book-entry securities accounts.[3] Book-entry account holders that own Treasury securities can house their securities directly in their accounts and can transfer the securities to other book-entry accounts by issuing appropriate instructions to the Federal Reserve.

Federal Reserve book-entry accounts are generally available only to depository institutions and certain other organizations, such as government-sponsored enterprises and foreign central banks. All other market participants establish ownership of Treasury securities through *commercial* book-entry accounts at depository institutions that act as custodians for their customers. Depository institutions that offer commercial

[3] See "Book-Entry Securities Account Maintenance and Transfer Services," Federal Reserve Banks Operating Circular no. 7, August 19, 2005. Garbade (2004) describes the origins of the Federal Reserve book-entry system.

book-entry accounts hold their customers' securities in their Federal Reserve book-entry accounts commingled with their own securities.

A market participant with a commercial book-entry account can transfer Treasury securities to another market participant through their respective custodians. For example, participant A can transfer a Treasury security to participant B by instructing its custodian to debit its commercial book-entry account and to transfer the security to B's custodian for credit to B's commercial book-entry account. Upon receipt of instructions, A's custodian will debit A's account and instruct the Federal Reserve to 1) debit its Federal Reserve book-entry account and 2) credit the Federal Reserve book-entry account of B's custodian. Following receipt of the security in its Federal Reserve book-entry account, B's custodian will complete the transfer by crediting B's commercial book-entry account. (If A and B have a common custodian, the transfer can be completed on the books of that common custodian without involving the Federal Reserve.)

2.2 Bilateral Settlement

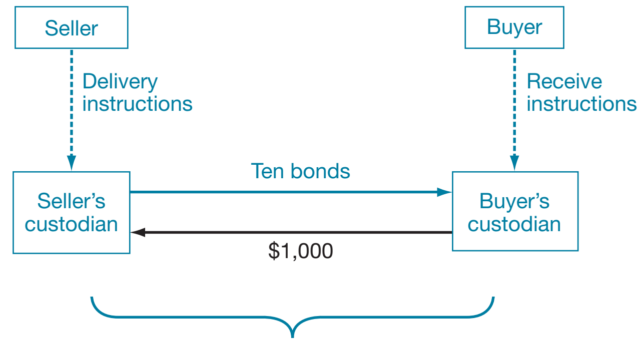
The simplest type of settlement occurs when a market participant has sold Treasury securities for bilateral settlement on a deliver-versus-payment basis. The sale may be a conventional sale of securities but it may alternatively be

The simplest type of settlement occurs when a market participant has sold Treasury securities for bilateral settlement on a deliver-versus-payment basis.

the starting leg, or the "off" leg, of a repurchase agreement. (We describe repurchase agreements in more detail below.)

Suppose, for example, an investor sells ten Treasury bonds at a price of $100 per bond for settlement on June 2. Following negotiation of the terms of the sale, the seller will instruct its custodian to send ten bonds to the buyer's custodian on June 2 against payment of $1,000. The buyer will concurrently instruct its custodian to receive, on June 2, ten bonds from the seller's custodian and to pay $1,000 upon receipt of the bonds. On June 2, the seller's custodian will instruct the Federal Reserve to 1) debit its Federal Reserve book-entry account for ten bonds, 2) credit the Federal Reserve book-entry account of the buyer's custodian for ten bonds *and simultaneously* debit the account of the buyer's custodian for the $1,000 due upon

Exhibit 1
Bilateral Settlement



Note: This transfer of bonds and funds is effected through the Federal Reserve if the seller and buyer have different custodians, and is effected on the books of the common custodian if they have the same custodian.

delivery, and 3) credit the seller's custodian's account for the $1,000. The resulting transfers of securities and funds are shown in Exhibit 1.[4]

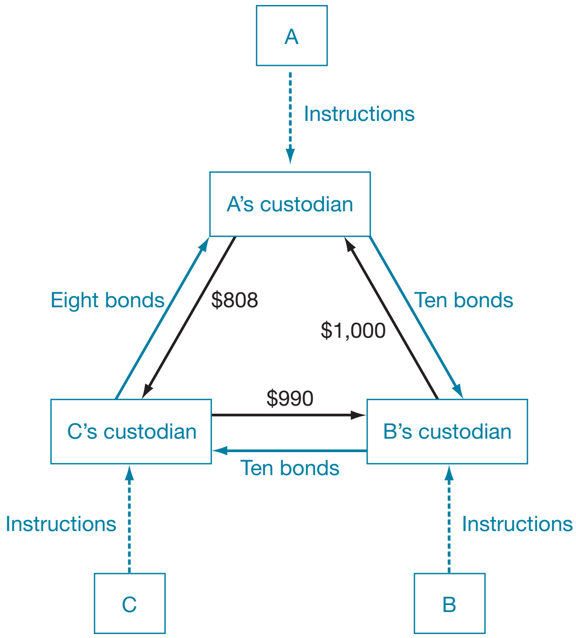
Following notification that ten bonds have come into its Federal Reserve book-entry account and that $1,000 has been withdrawn, the buyer's custodian will verify that the bonds and money are consistent with the buyer's instructions. In most cases, they are and the custodian will credit the buyer's account for the ten bonds and debit that account for $1,000. In some cases, however, the buyer will have provided different instructions—perhaps referencing a different security or a different invoice price—or no instructions. In any of these cases, the buyer's custodian will reverse the settlement, instructing the Federal Reserve to return the ten bonds and recover the $1,000 payment. The buyer and seller and their respective custodians will then have to communicate and come to a common understanding of the terms of the underlying transaction, following which the seller will reinitiate the settlement process.

2.3 Multilateral Net Settlement

Bilateral settlement is a simple process that satisfies the purpose of settlement: moving securities from sellers to buyers and moving funds from buyers to sellers. Alternative settlement structures, however, can sometimes be more efficient.

[4] In the event the buyer and seller have a common custodian, settlement can be completed on the books of the common custodian, with cash and securities moving between the accounts of the respective customers, without involving the Federal Reserve.

EXHIBIT 2
Bilateral Settlement of Three Transactions



Consider, for example, the case where:

- participant A sells ten bonds to participant B at a price of $100 per bond for settlement on the following business day,
- B sells ten of the *same* bonds to participant C at a price of $99 per bond, also for settlement on the following business day, and
- C sells eight of the *same* bonds to A at a price of $101 per bond, again for settlement on the following business day.

As shown in Exhibit 2, bilateral settlement of the three transactions requires the delivery of twenty-eight bonds against payments of $2,798.

As an alternative, the participants might agree to settle through a central counterparty (CCP). The CCP first marks all of the deliver and receive obligations to a common price—say, $100 per bond. After marking to the common price,

- A is obligated to deliver ten bonds to B against payment of $1,000,
- B is obligated to deliver ten bonds to C against payment of $1,000, and
- C is obligated to deliver eight bonds to A against payment of $800.

Marking to a common price results in gains for some participants and losses for others. In the example, B gains because it will receive more for the bonds sold to C than the original contract price and C loses for the same reason. These gains and losses are exactly offset with further agreements to make small side payments of cash. In particular:

- A agrees to pay $8 to the CCP,
  reflecting the $8 gain from marking the price of the eight bonds bought from C down from $101 per bond to $100 per bond,
- B agrees to pay $10 to the CCP,
  reflecting the $10 gain from marking the price of the ten bonds sold to C up from $99 to $100 per bond, and
- the CCP agrees to pay $18 to C,
  in compensation for the $8 loss from marking the price of the eight bonds sold to A down from $101 per bond, and for the $10 loss from marking the price of the ten bonds bought from B up from $99 per bond.

On the night before the settlement date, the CCP nets out the deliver and receive obligations of A, B, and C and novates[5] their respective contracts, becoming the buyer from every *net* seller and the seller to every *net* buyer, all at the common settlement price. After netting and novation:

- A is obligated to deliver two bonds to the CCP against payment of $200,
- B has no deliver or receive obligations, and
- the CCP is obligated to deliver two bonds to C against payment of $200.

On settlement day, the obligations of A to deliver two bonds to the CCP and the CCP to deliver two bonds to C are settled with bilateral deliver-versus-payment settlements. In addition, A, B, and the CCP make the agreed-upon side payments of cash. Exhibit 3 shows that multilateral net settlement requires the delivery of four bonds and payments of $436—about 15 percent of the deliveries and payments shown in Exhibit 2.

2.4 Some Concrete Identities

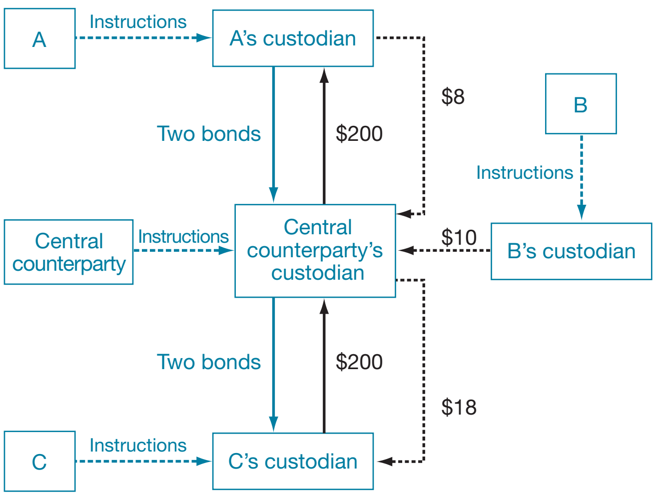
The foregoing description of settlement processes referred to abstract entities like "participant A" and an unnamed "central counterparty." Before we begin to discuss settlement fails, it may be helpful to identify some of the key participants in the Treasury market.

At the center of the market is a group of dealers that provide liquidity to customers, quoting bid prices at which they are willing to buy and offer prices at which they are prepared to sell. A subset of dealers, called "primary dealers," make markets

[5] To novate is to substitute one legal obligation for another.

Exhibit 3

Multilateral Net Settlement of Three Transactions



Note: Security settlements are shown with solid blue and black lines. Side payments, represented by dashed black lines, take place independently of security settlements.

to the Federal Reserve Bank of New York when the Bank is conducting open market operations on behalf of the Federal Reserve System.[6] Box 1 identifies the primary dealers as of mid-2008.

Dealers sometimes trade directly with each other, but more commonly through specialized interdealer brokers. A dealer that sells securities to another dealer through an interdealer broker agrees to deliver securities (against payment) to the broker. The broker, in turn, agrees to deliver the same securities (also against payment) to the ultimate buyer. This arrangement allows the dealers to trade on a "blind," or undisclosed, basis.

All of the dealers, and all of the interdealer brokers, maintain commercial book-entry accounts at one of two banks: JPMorgan Chase Bank, N.A., and The Bank of New York Mellon. These two "clearing" banks offer custodial services refined over many years to meet the needs of brokers and dealers that deliver and receive large volumes of securities on a daily basis.

The Fixed Income Clearing Corporation (FICC), a subsidiary of the Depository Trust & Clearing Corporation, is the central counterparty in the Treasury market. All of the primary dealers and all of the interdealer brokers, as well as a number of other market participants, are netting members of FICC. FICC maintains commercial book-entry accounts at both JPMorgan Chase and The Bank of New York Mellon and is prepared to receive securities from, and deliver securities to, any of its netting members in a timely and efficient fashion.

Beyond the dealers, the interdealer brokers, and FICC, the Treasury market consists of a large number of other participants, including "real-money" investors such as mutual funds, pension funds, and corporate treasurers, and "leveraged accounts" such as hedge funds. Some of these participants trade directly with dealers, others trade anonymously in electronic markets. All use custodians that offer more or less complex (and more or less costly) services tailored to their needs.

[6] See, generally, Federal Reserve Bank of New York, "Administration of Relationships with Primary Dealers," January 11, 2010, available at http://www.newyorkfed.org/markets/pridealers_ policies.html.

Box 1

Primary Dealers in Mid-2008[a]

Banc of America Securities LLC
Barclays Capital Inc.
Bear, Stearns & Co., Inc.[b]
BNP Paribas Securities Corp.
Cantor, Fitzgerald & Co.

Citigroup Global Markets, Inc.
Credit Suisse Securities (USA) LLC
Daiwa Securities America Inc.
Deutsche Bank Securities Inc.
Dresdner Kleinwort Securities LLC

Goldman, Sachs & Co.
Greenwich Capital Markets, Inc.
HSBC Securities (USA) Inc.
J. P. Morgan Securities Inc.
Lehman Brothers Inc.[c]

Merrill Lynch Government Securities Inc.[d]
Mizuho Securities USA Inc.
Morgan Stanley & Co. Incorporated
UBS Securities LLC

[a] Federal Reserve Bank of New York, "Primary Dealers List," July 15, 2008, available at http://www.newyorkfed.org/newsevents/news/markets/2008/an080715.html.

[b] Removed October 1, 2008, following its acquisition by J. P. Morgan Securities Inc.

[c] Removed September 22, 2008.

[d] Removed February 11, 2009, following its acquisition by Bank of America Corporation.

2.5 Settlement Fails

A settlement fail occurs when the obligation of a seller to deliver securities to a buyer remains outstanding following the close of business on the scheduled settlement date of a transaction. This can occur either because the seller's custodian failed to tender any securities to the buyer's custodian, or because the buyer's custodian rejected whatever securities were tendered by the seller's custodian. In the event of a settlement fail in Treasury securities, the market convention is to postpone settlement to the following business day without any change in the funds due upon delivery and (prior to May 2009) without any explicit penalty or charge.[7] The process of failing (to settle) and deferring settlement to the next business day can take place repeatedly, day after day, until settlement occurs or the trade is canceled.

A settlement fail occurs when the obligation of a seller to deliver securities to a buyer remains outstanding following the close of business on the scheduled settlement date of a transaction.

Settlement fails can occur for any of several reasons. First, a fail can result from miscommunication. A buyer and seller may not have a common understanding of the terms of a trade, or one or the other may have failed to communicate settlement instructions to its custodian, or may have communicated incorrect instructions, or one of the custodians may have misunderstood the instructions that it received. In any of these cases, the buyer's custodian will reject whatever securities are tendered by the seller's custodian. After becoming aware of the failed attempt to settle (or of the absence of any attempt to settle), the buyer and seller and their respective custodians communicate to resolve the problem. This usually results in successful settlement within a day or two.

A fail may also stem from operational problems. One well-known instance occurred on Thursday, November 21, 1985, when a computer outage at The Bank of New York (a predecessor of The Bank of New York Mellon) prevented that bank from effecting deliveries of Treasury securities. The bank was unable to resolve the problem until the following day, and had to finance overnight (at its own expense) the customer securities that it was unable to deliver. It borrowed in excess of $20 billion from the Federal Reserve Bank of New York and incurred interest expenses of $5 million.[8]

A settlement fail can also occur because the seller does not have the requisite securities in its commercial book-entry account. This is the most common reason for failing when fails are chronic, but it is usually avoided at other times by borrowing securities and delivering the borrowed securities.

[7] This convention was memorialized in Chapter 8, Section C, of the *Government Securities Manual* of the Public Securities Association: "If securities are not delivered on the agreed upon settlement date, there is a fail. Regardless of the date the securities were actually delivered, the buyer of the securities pays the seller the original settlement date figures." The Public Securities Association was the forerunner of the Bond Market Association, which joined with the Securities Industry Association in 2006 to form the Securities Industry and Financial Markets Association.

3. Repurchase Agreements and Borrowing Securities to Avoid or Cure Settlement Fails

A repurchase agreement is a sale of securities coupled with an agreement to repurchase the same securities at a specified price on a later date.[9] Market participants use repos to borrow money when they buy securities but do not have sufficient cash on hand to pay for them, that is, to finance long positions, as well as to borrow securities when they sell securities they do not already own, that is, to finance short positions.

A repo is analogous to a loan, where the proceeds of the initial sale correspond to the principal amount of the loan and the excess of the repurchase price over the original sale price corresponds to the interest paid on the loan. A market participant might, for example, sell securities for $10 million and simultaneously agree to repurchase the securities ten days later for $10,008,333. This is analogous to borrowing $10 million for ten days at an interest rate of 3 percent per annum.[10] Market participants commonly think of repos as loans, rather than as purchases and sales, and quote repos in terms of interest rates rather than in terms of sale and repurchase prices.[11]

[8] A Computer Snafu Snarls the Handling of Treasury Issues," *Wall Street Journal*, November 25, 1985, p. 58; Committee on Banking, Finance, and Urban Affairs (1985); "Fed is Queried on Failure of Bank Computer System," *Wall Street Journal*, December 13, 1985, p. 49; "Fed Weighs a Penalty," *New York Times*, December 13, 1985, p. D2; Sender (1986).

[9] Repurchase agreements are complex financial instruments whose contracting conventions have evolved over the past four decades. See Garbade (2006) and Fleming and Garbade (2003, 2004).

[10] $8,333 = (repo term of 10 days / 360 days in a year) × 3 percent per annum × $10 million, where the calculation uses the money market convention of a 360-day year.

[11] The quotation convention does not change the nature of a repo—a transaction in which one party sells securities subject to an agreement to repurchase them at a later date.

Repos are most commonly arranged on an overnight basis but can run for days or weeks. They can also be arranged on an "open," or continuing, basis (with a daily adjustment of the interest rate) at the mutual consent of the parties. Industry standard documentation for a repo provides that if the original seller fails to repurchase the securities on the agreed-upon repurchase date, the original buyer has the contractual right to, among other things, sell the securities to a third party and use the proceeds to satisfy the original seller's repurchase obligation. Conversely, if the original buyer does not deliver the securities back to the original seller on the repurchase date, the original seller has the contractual right to, among other things, use the funds that it otherwise would have used to repurchase the securities to "buy in," or replace, the securities.

3.1 Types of Repurchase Agreements

Repos come in two flavors: *general collateral* repos (used to borrow money) and *special collateral* repos (used to borrow securities).

General collateral repos: A general collateral repo is a repo in which the lender of funds is willing to accept any member of a stated class of securities as collateral. Any of a variety of securities is acceptable because the lender is concerned primarily with earning interest on its money and having possession of liquid assets that can be sold quickly in the event of a default by the borrower.

> *Repos come in two flavors:* general collateral *repos (used to borrow money) and* special collateral *repos (used to borrow securities).*

Interest rates on overnight general collateral repos are usually quite close to rates on overnight loans in the federal funds market. This reflects the essential character of a general collateral repo as a device for borrowing and lending money. Repo rates for the most liquid and creditworthy collateral, Treasury securities, are lowest. Repo rates for other classes of collateral, such as fixed-income securities issued by a federal agency or mortgage-backed securities issued by a government sponsored enterprise, are somewhat higher.

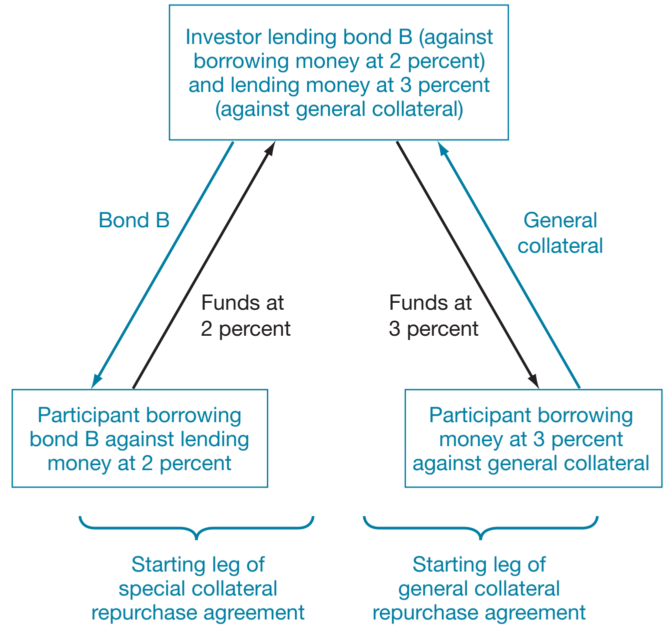
Special collateral repos: A special collateral repo is a repo in which the lender of funds designates a *particular* security as the *only* acceptable collateral.[12] Treasury market participants commonly lend money on special collateral repos in order to borrow specific securities that they need.

The interest rate on a special collateral repo is called a "specials rate." The owner of a Treasury security that other market participants want to borrow may be incentivized to lend the security if that owner is offered an opportunity to borrow money at a specials rate less than the Treasury general collateral repo rate. For example, if the rate on a special collateral repo involving bond B is 2 percent and the general collateral repo rate is 3 percent, an investor who owns bond B can earn a 100 basis point spread by lending the bond and borrowing money on a special collateral repo and then relending the money on a general collateral repo (Exhibit 4).

The difference between the general collateral repo rate for Treasury securities and the special collateral repo rate for a particular Treasury security is a measure of the "specialness" of the security and is commonly called the security's "specialness spread." We show below that a security's specialness spread is exactly the opportunity cost of borrowing the security to avoid or cure a settlement fail.

Exhibit 4

Lending Treasury Bond B (against Borrowing Money at 2 Percent) on a Special Collateral Repurchase Agreement and Relending the Money on a General Collateral Repurchase Agreement at 3 Percent



Note: For simplicity, the separate roles of custodians are not shown explicitly.

[12] See Duffie (1996), Keane (1996), Jordan and Jordan (1997), Fisher (2002), and Fleming and Garbade (2002).

3.2 Incentives, prior to May 2009, to Borrow Securities to Avoid or Cure a Settlement Fail

Prior to May 2009, sellers of Treasury securities, including short sellers, borrowed securities to avoid or cure settlement fails primarily because they did not get paid until they delivered the securities that they had sold. Prior to May 2009, market participants usually quantified the cost to a seller of a settlement fail in Treasury securities as the overnight Treasury general collateral repo rate—the rate the seller could have earned on a riskless overnight investment of the sale proceeds that it did not receive. (It should be noted, however, that even prior to May 2009, the cost of a settlement fail was not limited to foregone interest earnings. Settlement fails also expose market participants to the risk of counterparty insolvency and can lead to increased capital charges for some participants. These other costs are discussed in Box 2.)

Prior to May 2009, a seller had an incentive to avoid failing to deliver a security (by borrowing the security on a special collateral repo and delivering the borrowed security) as long as the cost of borrowing the security was less than the cost of failing.

A seller who does not have the securities needed to settle a sale can avoid failing by borrowing (on a special collateral repo) the securities that it needs. However, borrowing securities is not costless because the borrower has to lend money (on the special collateral repo) at a rate lower than the general collateral repo rate that it could have earned on the money. The cost of borrowing securities to avoid a fail in Treasury securities may be quantified as the difference between the overnight Treasury general collateral repo rate (the rate the borrower could have earned on its money) and the overnight special collateral repo rate on the borrowed securities (the rate the borrower actually earns on its money)—that is, the securities' specialness spread.

Prior to May 2009, a seller had an incentive to avoid failing to deliver a security (by borrowing the security on a special collateral repo and delivering the borrowed security) as long as the cost of borrowing the security was less than the cost of failing. This was certainly the case if the specialness spread for the security was less than the general collateral repo rate or, equivalently, if the special collateral repo rate for the security was greater than zero.[13] As long as a seller could earn more than a *de minimis* amount of interest on a special collateral repo, it made economic sense to lend the money, earn the interest, and avoid the fail.

Box 2

Other Costs of Settlement Fails

Settlement fails impose risks on both buyers and sellers and can lead to increased capital charges for some market participants.

A buyer who fails to receive securities faces the risk that the seller might become insolvent before the transaction is settled and that, to replace the securities, it will have to pay more than the price negotiated with the insolvent seller. Conversely, a seller who fails to deliver securities faces the risk that the buyer might become insolvent before the transaction is settled and that the seller will then have to sell its securities to someone else at a price lower than the price negotiated with the insolvent buyer. These risks may be small for fails that are no more than a day or two old (because securities prices usually do not change much from day to day), but they can become significant when fails persist for weeks or months.

In view of the greater risks faced by both buyers and sellers during the life of a long-outstanding, or "aged," fail, some market participants are required to absorb incremental capital charges when a settlement fail has been outstanding for more than a few weeks.[a] Such capital charges can drain capital from more rewarding activities and limit balance sheet capacity, thereby imposing opportunity costs on market participants—and if sufficiently widespread and chronic can threaten overall market functioning.[b]

[a] See, for example, the net capital rule of the Securities and Exchange Commission, Code of Federal Regulations, Chapter 17, Section 15c3-1.

[b] See, for example, "Minutes of the Meeting of the Treasury Borrowing Advisory Committee of the Bond Market Association, November 4, 2003," November 5, 2003, available at http://www.ustreas.gov/press/releases/js933.htm ("While the situation is much improved since this past summer, members commented that fails were still at an elevated level which does hurt general market liquidity because dealers are forced to reduce their market making activities as the fails take up space on their balance sheets."). One Treasury official suggested that opportunity costs resulting from higher capital charges might not be all bad. See "Remarks by Jeff Huther, Director of the Office of Debt Management, to the Bond Market Association's Annual Meeting," April 22, 2004, available at http://www.treas.gov/press/releases/js1455.htm (noting that "capital charges resulting from chronic—widespread and persistent—fails soak up dealer capital that might otherwise be used to support profit-making activities, thereby focusing management attention on the underlying fails problem and incentivizing managers to remedy the situation.").

[13] Using economic terminology, let R_{gc} denote the general collateral repo rate and R_{sp} denote the special collateral repo rate for the security. The specialness spread $R_{gc} - R_{sp}$ will be less then R_{gc} if and only if R_{sp} is greater than zero.

3.3 Equilibrium in the Market for Special Collateral Repos

The market for borrowing and lending a particular Treasury security comes into equilibrium as a result of fluctuations in the special collateral repo rate for the security *relative to* the Treasury general collateral repo rate. If the demand to borrow the security is modest relative to the supply available for lending, a market participant seeking to borrow the security will usually be able to lend its money at a specials rate no lower than about 15 to 25 basis points below the general collateral repo rate. However, if the demand to borrow the security expands, some borrowers (in order to avoid failing) will be willing to accept less interest on the money they lend. Downward pressure on the specials rate for the security relative to the general collateral repo rate makes lending the security more remunerative, thereby attracting additional lenders. It may also ration some borrowers out of the market, particularly short sellers who decide to liquidate their short positions rather than continue to finance those positions on special collateral repos earning lower rates of interest. The collateral market will return to equilibrium, that is, to a state where the quantity of the security sought to be borrowed at the prevailing specials rate equals the quantity of the security available for lending at that rate, when the lower specials rate has attracted enough new lenders and/or rationed enough borrowers out of the market.

4. Federal Reserve and Industry Efforts to Mitigate Settlement Fails

Treasury market participants have an interest in mitigating settlement fails in order to limit their net interest expenses as well as their exposure to the risk of counterparty insolvency—a risk explained in Box 2. The Federal Reserve has a separate interest in mitigating settlement fails to maintain the liquidity and efficiency of the market in which it conducts open market operations. (A high volume of fails can lead market participants to reduce, or even withdraw from, their normal activities. Such activities include dealers making markets for customers, investors lending securities to dealers to facilitate settlement of dealer sales, and arbitrageurs seeking to exploit, and thereby eliminate, price relationships that present abnormal profit opportunities.)

Since 1969, the Federal Reserve has sought to mitigate settlement fails by lending Treasury securities to primary dealers to facilitate settlement of dealer sales.[14] (However, the Federal Reserve lends against collateral, rather than cash, to insulate the supply of reserves available to the banking system from securities lending operations.) Pursuant to the terms and conditions of the lending program in effect in mid-2008,[15] each business day at noon the Federal Reserve Bank of New York offered to lend on an overnight basis up to 90 percent of the amount of each Treasury security beneficially owned in the Federal Reserve's System Open Market Account (SOMA), subject to an upper limit of the amount of an issue actually in the account.[16] Primary dealers bid for a security by specifying the quantity desired (in increments of $1 million) and the overnight loan fee they were willing to pay, expressed as a rate per annum, subject to a minimum fee of 50 basis points.[17] Bidding closed at 12:15 p.m. and loans were awarded to the highest bidders at their bid rates[18] until all of the securities available for lending were allocated or all of the bidders had been satisfied.[19] During the first six months of 2008, the Federal Reserve Bank of New York lent an average of $12.2 billion of securities per day, distributed over an average of twenty-three different issues, at an average loan fee of 61 basis points per annum.

Treasury market participants have an interest in mitigating settlement fails in order to limit their net interest expenses as well as their exposure to the risk of counterparty insolvency.

[14] In authorizing the loan of Treasury securities from the System Open Market Account in 1969, the Federal Open Market Committee stated that the action "was taken after the Manager [of the System Open Market Account] had advised that the problem of delivery failures in the Government securities market had worsened significantly over the past year, partly because private facilities for lending such securities had become inadequate; that delivery failures were markedly impairing the performance of the market; and that the functioning of the market would be improved if securities held in the System Open Market Account could be lent, for the express purpose of avoiding delivery failures, to Government securities dealers doing business with the Federal Reserve Bank of New York . . . ." (Federal Reserve *Bulletin*, January 1970, p. 32).

[15] "SOMA Securities Lending Program Terms and Conditions (Revised)," Federal Reserve Bank of New York, August 22, 2008, available at http://www.newyorkfed.org/newsevetns/news/ markets/2008/an082208.html. Details of the history of the Fed's lending program appear in Fleming and Garbade (2007).

[16] In order to avoid failing itself, the Fed does not agree to lend securities unless it has actual possession of the securities at the time of an auction. Thus, it will not agree to lend securities that it lent the preceding business day and that have not yet been returned.

[17] A loan fee for a security is approximately equivalent to the security's specialness spread. See Fleming and Garbade (2007). The minimum fee avoids crowding out private lenders when security loan markets are functioning normally, but it has been reduced to nearly zero when those markets are not functioning well.

[18] The auction for each security is a discriminating, or multiple-price, auction.

Treasury market participants have also acted cooperatively to mitigate settlement fails and otherwise reduce the cost of settling transactions. Between 1986 and 1988, dealers in Treasury securities organized the Government Securities Clearing Corporation (GSCC) to serve as a central counterparty in interdealer transactions in Treasury and related securities. As explained earlier, multilateral net settlement through a central counterparty economizes on the quantity of securities that have to be delivered to settle a given volume of transactions. GSCC also implemented a trade confirmation protocol that essentially eliminated interdealer fails due to miscommunication between dealers, as well as a procedure for marking failed trades to current market prices that materially reduced the consequences of counterparty insolvency. The GSCC extended its net settlement system to include Treasury auction takedowns in 1994 and repurchase agreements in 1995.[20] In 2002, GSCC became a wholly owned subsidiary of Depository Trust & Clearing Corporation and was renamed the Fixed Income Clearing Corporation.

5. Chronic Settlement Fails

Demand to borrow a security (relative to the supply available for lending) can sometimes be large enough to drive the specials rate for the security down to near zero. Prior to May 2009, sellers would then become largely indifferent between a) failing and b) borrowing the security to avoid failing. In this extreme case, any unsatisfied demand to borrow would spill over into fails. Fails could expand further if security lenders, observing a growing incidence of settlement fails, declined to continue lending out of concern that their securities may not be returned on a timely basis.[21] Fails could expand still further if, as explained in Box 3, market participants concluded that they could acquire a cheap option on a future increase in a specials rate by contracting to sell a security in a special collateral repo and then strategically failing to deliver the security. More generally, settlement fails could become *chronic* when the specials rate for a security was driven down to near zero.

Three episodes of chronic fails have been described in the literature: in May and June of 1986,[22] following the terrorist attack on the World Trade Center on September 11, 2001,[23] and during the summer of 2003.[24]

Box 3
Strategic Fails

When the specials rate for a security is close to zero in a market without a fails charge convention, a market participant with no position in the security may sometimes agree to lend the security on a term repurchase agreement and then fail, intentionally, on the starting leg of the repo.

Suppose, for example, the three-week specials rate for a five-year note is 10 basis points and that XYZ Co. believes the 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 $2,917 at the end of three weeks.[a] It will owe this amount even if it fails to deliver the note any time during the three-week interval.

XYZ Co. has effectively purchased (for $2,917, payable in three weeks) an option 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 could choose to let its option expire unexercised and simply pay the $2,917 premium at the end of three weeks. However, if XYZ Co.'s expectations prove correct, it can exercise the option after one week by borrowing the five-year note for two weeks against lending $50 million at 50 basis points (earning interest of $9,722[b]) and delivering the note in (delayed) settlement of its earlier negotiated three-week repurchase agreement. The $50 million received from delivering the note funds the loan that allows XYZ Co. to borrow the note, and XYZ Co. has net interest earnings of $6,805 ($6,805 = $9,722 interest income, less $2,917 interest expense).

More generally, a very low specials rate presents an opportunity to speculate on an increase in the rate—by lending on, and then failing on, a special collateral repurchase agreement—with limited downside exposure.[c] In the limit, a repo rate of zero may be viewed by some participants as a "risk-free" opportunity to intentionally fail and either profit or break even. Such practices can lead to an increase in aggregate settlement fails and the associated indirect costs discussed in Box 2.

[a] $2,917 = (21/360) × 0.10 percent per annum × $50 million.

[b] $9,722 = (14/360) × 0.50 percent per annum × $50 million.

[c] Strategic fails are noted in Fleming and Garbade (2002, p. 47), Fleming and Garbade (2004, pp. 3-4), and in "Remarks by Jeff Huther, Director of the Office of Debt Management to the Bond Market Association's Annual Meeting," April 22, 2004, available at http://www.treas.gov/press/releases/js1455.htm.

[19] Awards were subject to three limitations: 1) no dealer could have outstanding borrowings of more than 25 percent of the amount of an issue beneficially owned in the System Open Market Account, 2) no dealer could have outstanding borrowings of more than $750 million of any single issue, and 3) no dealer could have outstanding borrowings of more than $3 billion of securities in total.

[20] See Garbade and Ingber (2005) and Ingber (2006).

[21] This can quickly lead to a self-reinforcing and destabilizing cycle, with lenders withdrawing collateral out of a concern that borrowers may fail to return the securities, thereby increasing the incidence of settlement fails and triggering further collateral withdrawals.

[22] Cornell and Shapiro (1989). Cornell and Shapiro do not discuss fails directly, but do document a near-zero specials rate for the 9 1/4 percent Treasury bond of February 2016 and discuss the reasons for that low rate. The existence of widespread settlement fails in the issue was common knowledge among market participants at the time.

[23] Fleming and Garbade (2002).

[24] Fleming and Garbade (2004). See also Fleming and Garbade (2005).

5.1 The 1986 Episode

An "on-the-run" Treasury security is the most recently auctioned security in a given series, such as the most recently auctioned six-month bill or ten-year note. In late April 1986, dealers began to sell the on-the-run thirty-year Treasury bond (the 9 1/4 percent bond of February 2016) short in anticipation of bidding for the new thirty-year bond that would be announced on April 30 and auctioned on May 8 for settlement on May 15.[25] Such short selling in advance of an auction announcement was normal and customary dealer behavior.[26] However, dealers soon found themselves unable to borrow enough 9 1/4 percent bonds to finance their short positions, in part because a significant quantity of the bonds was owned by investors who declined to lend.[27] Strong dealer demand and limited supply combined to drive the special collateral repo rate for the bonds down to about 5 basis points, and dealers began to fail on their settlement obligations. Failing, however, was expensive because the Treasury general collateral repo rate was about 6.75 percent, so dealers with short positions had an economic incentive to cure their fails another way: by buying (rather than borrowing) the 9 1/4 percent Treasury bonds and delivering the purchased bonds—thereby closing out their short positions. They bid up the price of the 9 1/4 percent bonds, relative to the prices of other Treasury bonds with similar maturities, until the higher price induced holders to sell the 9 1/4 percent bonds and replace them with higher yielding substitutes, thereby allowing dealers to cover their short positions.[28]

[25] Department of the Treasury, Securities and Exchange Commission, and Board of Governors of the Federal Reserve System (1992, p. B-1, footnote 1) ("participants sold the outstanding 9 1/4 percent bond . . . to prepare for the roll into the WI ["when-issued"] thirty-year bond."). Following the announcement of the forthcoming auction, dealers planned to buy the on-the-run thirty-year bond (thereby covering their previous short sales of that bond) against selling the WI thirty-year bond short. The transactions would leave them with short positions in the WI bond that they could cover in the auction.

[26] Department of the Treasury, Securities and Exchange Commission, and Board of Governors of the Federal Reserve System (1992, p. 10, footnote 11) ("dealers . . . sold [the 9 1/4 percent bond] short as part of a trading strategy that had worked in the past as they prepared to bid for a new thirty-year bond."). The sequence of shorting the on-the-run thirty-year, rolling the short into the WI thirty-year, and then bidding to buy the WI thirty-year in the auction was part of the process whereby dealers distributed new bonds to market participants.

[27] Department of the Treasury, Securities and Exchange Commission, and Board of Governors of the Federal Reserve System (1992, p. 10, footnote 11; p. B-1, footnote 1) ("some institutional investors did not make the [9 1/4 percent bonds] available to the repo market" and "securities needed to [finance] short positions were not readily available to the repo market.").

[28] Cornell and Shapiro (1989, pp. 303-4) suggest that the 9 1/4 percent bond of February 2016 was overvalued by as much as 7 percent of principal value compared with one close substitute (the 9 7/8 percent bond of November 2015).

5.2 The 2001 Episode

The 2001 fails episode was attributable, in the first instance, to the terrorist attack on the World Trade Center on September 11, 2001. The attack destroyed the offices of several interdealer brokers and impaired telecommunication services throughout lower Manhattan. GSCC recorded $266 billion in interdealer settlement fails on September 11 and $440 billion in interdealer fails on September 12.[29] Sellers tried to borrow the securities needed to cure their fails but holders realized that, in view of the severe operational problems, their securities might not be returned on a timely basis and they consequently declined to lend.[30] The contraction in the supply of collateral pushed specials rates to near zero and settlement fails remained elevated. Daily average fails in Treasury securities reported by primary dealers to the Federal Reserve[31] reached $200 billion per day during the week of September 13-19 and continued high through early October. Settlement fails were particularly high for the on-the-run five-year note (the 4 5/8 percent note

Settlement fails began to shrink to more normal levels after the Treasury reopened the on-the-run ten-year note in an extraordinary unscheduled auction offering on Thursday, October 4, and after officials indicated that they might reopen the on-the-run five-year note as well.

[29] Fleming and Garbade (2002, p. 46).

[30] See "After Attack, Settlement Woes Still Clogging Repo Market," *Dow Jones Newswires*, September 26, 2001, 9:05 (noting "a general reluctance among large portfolios to lend their securities" and "in a chain reaction, the fear of failing trades is 'causing portfolio managers, securities lending desks and foreign central banks to hold even tighter on to their collateral,' which is exacerbating the situation . . . ."); "Treasury Market is Faced with Incomplete Trades," *Wall Street Journal*, October 3, 2001, p. C10 ("Because of the rate of fails . . . dealers are reluctant to use their securities as collateral. They are worried that they might not have securities delivered to them . . . . "); "U.S. Sells $6 Billion in 10-Year Notes to Help Overcome Shortage," *Bloomberg News*, October 4, 2001, 13:16 (quoting Peter Fisher, Under Secretary of the Treasury for Domestic Finance, as saying that "the cause of the fails is [in part] the result of . . . reluctance by institutional investors to lend into a market that is suffering from extraordinarily high fails levels."); and "U.S. Acts on Shortage of Treasuries," *New York Times*, October 5, 2001, p. C1 ("With the prospect that securities might not be returned, both dealers and large investors have become unwilling to lend them in the repo market.").

[31] Fleming and Garbade (2005) describe the settlement fails data reported by primary dealers to the Federal Reserve. Unless otherwise noted, this article measures settlement fails as the daily average over weekly intervals of the average of cumulative primary dealer fails to receive Treasury securities during a week and primary dealer fails to deliver Treasury securities over the same week. The Federal Reserve does not publish data on settlement fails on a day-by-day basis.

of May 2006) and the on-the-run ten-year note (the 5 percent note of August 2011).

Settlement fails began to shrink to more normal levels after the Treasury reopened the on-the-run ten-year note in an extraordinary unscheduled auction offering on Thursday, October 4, and after officials indicated that they might reopen the on-the-run five-year note as well.[32] Peter Fisher, the Under Secretary of the Treasury for Domestic Finance, stated that the Treasury reopened the ten-year note "to reduce the risk that . . . settlement problems turn into a much bigger problem for the Treasury market . . . ."[33] Fisher went on to observe that "we have something that is self-compounding. There is some point at which your fails pile up, and that is the point at which you damage the price-discovery process and the smooth operating of the Treasury market."

The actions of Treasury officials convinced market participants that the Treasury would take unprecedented steps to facilitate settlements and maintain market liquidity. Holders of the on-the-run five- and ten-year notes began to make the notes available, and the level of fails subsided.[34]

5.3 The 2003 Episode

The 2003 fails episode was attributable, in the first instance, to a heavy volume of short sales of the on-the-run ten-year note (the 3 5/8 percent note of May 2013) in late June 2003 by market participants seeking to hedge their interest rate risk on long positions in other fixed-income securities.[35] The short sales created an unusually large demand to borrow the note that drove the specials rate for the note down to zero, after which the residual, unsatisfied demand spilled over into fails. The fails became chronic when investors began to withdraw from lending the note. Daily average fails in Treasury securities reported by primary dealers to the Federal Reserve went from $25 billion per day during the week ending June 18 to $103 billion per day during the week ending July 2, and topped out at $232 billion per day during the week ending August 20. Settlement fails persisted for months[36] and were not fully resolved until the end of the year, following an offering of a new series of ten-year notes in November.[37]

5.4 Proposals to Mitigate Chronic Settlement Fails

The 2003 episode had a strong impact on the thinking of market participants. Unlike the 1986 episode, which was short-lived and quickly forgotten, and unlike the 2001 episode, which clearly stemmed from unusual circumstances, the 2003 episode was lengthy, large-scale, and stemmed from a marketplace activity—hedging—that was a very ordinary occurrence. The 2003 episode raised the question of whether something should be done, by government officials or by private sector market participants, to mitigate chronic fails.

> *The 2003 [fails] episode raised the question of whether something should be done, by government officials or by private sector market participants, to mitigate chronic fails.*

The key difference between the 1986 and 2003 episodes was the level of the Treasury general collateral repo rate. In May 1986, the overnight general collateral repo rate was about 6.75 percent. That made it costly to continue to fail even after the special collateral repo rate on the 9 1/4 percent bonds of February 2016 had been driven down to near zero and the economic incentive to avoid failing by borrowing the bonds had been eliminated. The high cost of failing incentivized short sellers to cover their short positions with outright purchases, and they bid up the price of the 9 1/4 percent bonds to a level that gave holders an economic incentive to swap out of the issue and into higher yielding substitutes.

In mid-2003, however, the overnight Treasury general collateral repo rate was about 1 percent, so the cost of failing was modest. Short sellers had little incentive to cover their

[32] "U.S. Sells $6 Billion in 10-Year Notes to Help Overcome Shortage," *Bloomberg News*, October 4, 2001, 13:16, and "U.S. Acts on Shortage of Treasuries," *New York Times*, October 5, 2001, p. C1 (both quoting the Under Secretary of the Treasury for Domestic Finance as saying that the Treasury might reopen the five-year note in the next week).

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

[34] See, for example, "Remedial Reopenings and the Treasury Supply Outlook," *Money Market Observer*, October 8, 2001 ("Dealers reported a dramatic reduction in the volume of fails on [October 5] after the settlement of the additional $6 billion of ten-year notes . . . ."), and "Another Emergency Treasurys Sale Looks Unlikely as Shortages that Hamstrung 'Repo Market' Ease," *Wall Street Journal*, October 11, 2001, p. C17.

[35] "Supply Dries Up Following Fall in Prices," *Financial Times*, August 23, 2003, p. 27 (reporting that "Demand for Treasuries from some quarters has also risen as prices have fallen because many institutions want to borrow the securities and 'short' them in the expectation that prices will continue to drop. Traders say hedged positions for the [on-the-run ten-year note] now exceed the amount of Treasury securities available.").

[36] See, for example, "Treasury Issue Remains a Headache," *Wall Street Journal*, November 17, 2003, p. C13.

[37] "California Standoff Dims Prospects," *Wall Street Journal*, December 9, 2003, p. C17 (reporting "progress for the . . . May ten-year note, which traders said appeared to be emerging from six months of gridlock, thanks to supply that entered the market last week. The note was trading in positive territory in the repurchase-agreement market. For months, it had been stuck at 0% . . . .").

short positions with outright purchases after demand to borrow the May 2013 ten-year note had driven the specials rate on the note down to near zero and eliminated the incentive to borrow the note to avoid failing.

Market participants and government officials learned from the 2003 episode that settlement fails were liable to become chronic quickly when short-term interest rates are low,[38] and they began to contemplate institutional innovations to avoid, or at least mitigate, chronic settlement fails. Most discussions centered around three possibilities:

- a regular program to reopen an issue when settlement fails in the issue become chronic,
- a securities lending facility run by the Treasury Department, and
- a fee to be paid by failing sellers to their counterparties to incentivize the sellers to resolve their fails.

Reopenings: Reopening an issue to alleviate chronic fails was exactly what the Treasury did when it reopened the on-the-run ten-year note on October 4, 2001.[39] However, Treasury officials were reluctant to institutionalize reopenings as a device to mitigate chronic fails. Three months after the 2001 reopening, Under Secretary Fisher told market participants that while "it would be imprudent of me to say that the Treasury will never again hold such an auction . . . you should not count on it, you should not expect it . . . ."[40] The problem was that reopenings in response to chronic fails ran counter to "regular and predictable" issuance, a cornerstone of Treasury debt management since the 1970s.[41] Treasury officials were concerned that the uncertainties engendered by an unpredictable reopening program would raise borrowing costs over the long run.[42]

A Treasury lending facility: Like reopenings, a *Treasury* lending facility would involve additional issuance from the Treasury. Unlike reopenings, a Treasury *lending* facility would increase supply on only a temporary basis. Such a facility was put forth as a "straw man" in a Treasury white paper published

Market participants and government officials learned from the 2003 episode that settlement fails were liable to become chronic quickly when short-term interest rates are low.

in May 2006.[43] The white paper was written to stimulate public discussion of mechanisms to make available "an additional, temporary supply of Treasury securities on rare occasions when market shortages threaten to impair the functioning of the market for Treasury securities and broader financial markets . . ."

However, Treasury officials questioned whether the Secretary of the Treasury has statutory authority to issue securities on a temporary basis to alleviate chronic settlement fails. Federal law provides that "the Secretary of the Treasury may borrow on the credit of the United States Government amounts necessary for expenditures authorized by law and may issue bonds of the Government for the amounts borrowed."[44] Similar provisions authorize the issuance of notes and bills.[45] The 2006 Treasury white paper suggested that "the Treasury would likely need to pursue new authority to issue securities for the purpose of securities lending. . . ."[46]

[38] See, for example, "Remarks by Jeff Huther, Director of the Office of Debt Management, to the Bond Market Association's Annual Meeting," Department of the Treasury, April 22, 2004, available at http://www.treas.gov/press/releases/js1455.htm ("The heart of the pricing problem last year was, unquestionably, the low federal funds rate and the consequently low ceiling on the cost of financing a short position."); Department of the Treasury (2006, p. 26, p. 174, footnote 2) ("The potential for chronic fails episodes thus increases in a very low interest rate environment such as that prevailing during the summer of 2003."); and "Statement of Under Secretary for Domestic Finance Randal K. Quarles to Bond Market Association Annual Meeting," Department of the Treasury, May 19, 2006, available at http://www.treas.gov/press/releases/js4274.htm ("When the central bank wishes to establish very low short-term rates, the maximum degree of specialness will be quite small. During these periods, we might expect to see greater incidence of fails episodes because the cost of failing is low.").

[39] That reopening was not the first time the Treasury increased the supply of a security in response to unusual market conditions. On November 3, 1992, the Treasury announced that it would reopen the 6 3/8 percent note of August 2002, at that time the on-the-run ten-year note, "in order to alleviate an acute, protracted shortage" of the note. See "Treasury November Quarterly Financing," Office of Financing, Department of the Treasury, November 3, 1992.

[40] "Remarks by Peter R. Fisher, Under Secretary of the Treasury for Domestic Finance, Before the Bond Market Association Legal and Compliance Conference," January 8, 2002, available at http://www.treas.gov/press/releases/po.906.htm.

[41] Garbade (2007) describes the emergence of "regular and predictable" as a Treasury debt management strategy.

[42] See "Remarks of Undersecretary of the Treasury Peter R. Fisher to the Futures Industry Association, Boca Raton, Florida," March 14, 2002, available at http://www.treas.gov/press/releases/po1098.htm.

[43] Department of the Treasury (2006). The public responses to the white paper are available at http://www.ustreas.gov/offices/domestic-finance/debt-management/slf-comments.pdf. See also Garbade and Kambhu (2005). A Treasury lending facility was also recommended by the Treasury Borrowing Advisory Committee following the chronic fails of late September and early October 2001; see "Report to the Secretary of the Treasury from the Treasury Advisory Committee of the Bond Market Association," October 30, 2001, available at http://www.ustreas.gov/offices/domestic-finance/debt-management/adv-com/reports/rpt-2001-q4.pdf ("members overwhelmingly felt that Treasury should expand their ability to enhance liquidity in the Treasury market. To accomplish this, they could set up a repo facility to help alleviate protracted shortages, in particular, large and persistent fails . . . .").

[44] 31 U.S.C. § 3102 (2010).

[45] See 31 U.S.C. § 3103-3104 (2010).

[46] Department of the Treasury (2006, pp. 26, 178).

A fails charge: In 2002, two economists at the Federal Reserve Bank of New York suggested that "chronic fails can also be alleviated by increasing the cost of failing with a penalty fee."[47] The economists noted that such a fee would give sellers an economic incentive to borrow securities to avoid failing even when the special collateral repo rate for the securities was close to zero. They further noted that a fails charge might lead market participants to borrow securities against lending money at *negative* specials rates (in order to avoid the fails charge) and that such negative specials rates could attract additional securities lenders (because they would receive, rather than pay, interest on the money they borrowed against lending securities).

The economists suggested that a fails charge might be set at some threshold rate minus the general collateral repo rate, with a minimum of zero.[48] The fails charge would be above zero only if the general collateral repo rate was below the threshold rate and would not be higher than what was necessary to bring the total cost of failing to the threshold rate. (For example, if the threshold rate is 5 percent and the general collateral repo rate is 3 percent, the fails charge would be 2 percent and the total cost of failing would be 5 percent.) They further suggested that the fails charge could be instituted through a "good-practice" recommendation of the Bond Market Association.[49] The economists noted that a fails charge could be implemented implicitly by reducing the invoice price on a transaction each day the seller fails—a material departure from the existing convention of deferring settlement at an unchanged invoice price—but observed that "the operational burden of changing an invoice price following a delay in settlement would undoubtedly be substantial."[50]

5.5 Inaction prior to the Insolvency of Lehman

Following the 2003 episode of chronic settlement fails, both government officials and private sector market participants understood that chronic fails are prone to blossom in an environment of low interest rates. Several parties had identified ways to address the problem, but each of the suggestions had a material deficiency. Treasury officials asked private sector participants to address the problem, but nothing substantive came of their requests.[51] No significant progress was made with respect to addressing the problem of chronic fails before the insolvency of Lehman in the fall of 2008.

[47] Fleming and Garbade (2002, p. 52).

[48] Fleming and Garbade (2002, p. 53).

[49] The Bond Market Association joined with the Securities Industry Association in 2006 to form the Securities Industry and Financial Markets Association.

[50] Fleming and Garbade (2002, p. 52).

6. Chronic Settlement Fails in the Wake of the Insolvency of Lehman

The announcement, early in the morning of Monday, September 15, 2008, that Lehman was insolvent triggered a "flight to safety" that, by the close of trading that day, pushed the yield on four-week Treasury bills down to 36 basis points, 100 basis points lower than the yield on the preceding Friday. Yields on longer term bills also moved sharply lower. By the close of trading on Wednesday, September 17, yields on four-week bills were down to 7 basis points. Over the balance of the month, four-week-bill yields fluctuated between about 10 basis points and 100 basis points—well below the 1.50 to 1.85 percent range that had prevailed since the beginning of August (Chart 1).

> *The first response to the rising tide of settlement fails was the decision of the Federal Reserve to relax the terms of its securities lending program.*

Greater demand for high-quality, short-term debt also drove down repo rates on Treasury collateral. The overnight Treasury general collateral repo rate averaged 90 basis points between September 15 and September 30, well below the 2 percent level that had prevailed during the preceding six weeks (Chart 2).

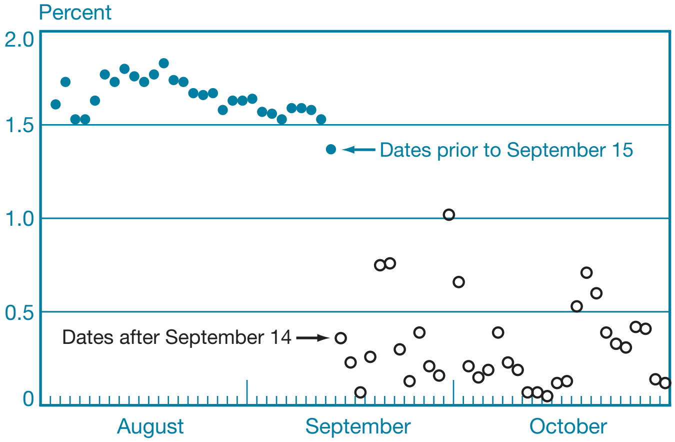
In the wake of Lehman's insolvency and in the midst of the ensuing flight to safety, investors became increasingly reluctant to lend Treasury securities.[52] Unable to replace their maturing borrowings, dealers began to fail on their delivery obligations.

[51] "Minutes of the Meeting of the Treasury Borrowing Advisory Committee," November 4, 2008, available at http://www.treas.gov/press/releases/hp1239.htm (stating that "Since November 2003, Treasury has repeatedly asked the private sector to address [the fails] issue proactively. On several occasions, market participants have emphatically stated that they would resolve the situation without government intervention, but such steps have not been implemented."). See also Wrightson, *Federal Reserve Data*, October 17, 2008 (stating that "the repo market has managed to fend off regulatory reform in past cycles."), and "The Treasury Market Reaches Breaking Point," *Euromoney*, December 1, 2008 (quoting a former Treasury employee as saying that "It was politically difficult to convince the market to put a stop to fails to deliver in treasuries. There were some forceful voices insisting that if the Treasury got involved, they would take the incentives out of the specials market altogether. Those making their living as specialist dealers, as well as those making a living shorting securities outright, were worried about potential supply changes which would eliminate trading opportunities for them.").

Chart 1

Yields on Four-Week Treasury Bills

August to October 2008

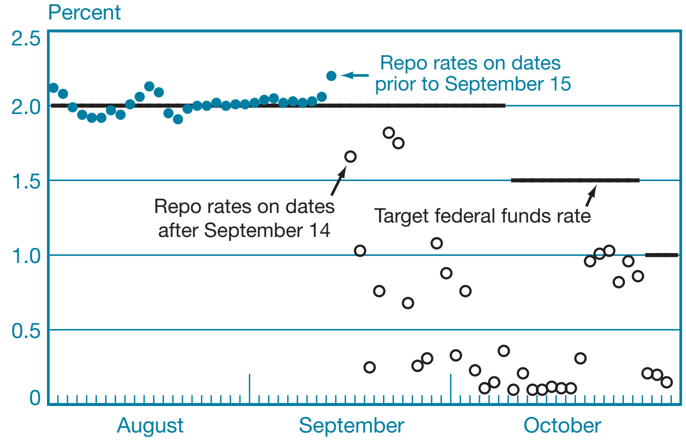


Source: Federal Reserve Statistical Release H.15.

Chart 2

Target Federal Funds Rates and Rates on Overnight General Collateral Repurchase Agreements on Treasury Securities

August to October 2008



Sources: Federal Reserve Statistical Release H.15; Board of Governors of the Federal Reserve System.

The fails persisted because the low general collateral repo rate left sellers with little incentive to cure the fails. Primary dealer settlement fails in Treasury securities mushroomed to an average of $253 billion per day during the week of Thursday, September 18, to Wednesday, September 24—far in excess of the level that had prevailed in August and the first half of September (Chart 3). And unlike earlier episodes, fails in the wake of Lehman's insolvency were not concentrated in one or two issues; rather, they involved securities across the entire yield curve.

The first response to the rising tide of settlement fails was the decision of the Federal Reserve to relax the terms of its securities lending program. As shown in Table 1, on Tuesday, September 23, the Fed raised the limit on total borrowings by a single dealer from $3 billion to $4 billion. Loans to primary dealers from the SOMA portfolio reached new heights but primary dealer settlement fails continued to rise, averaging $342 billion per day over the interval from September 25 to October 8 (Chart 3).

On Wednesday, October 8, both the Federal Reserve and the Treasury acted in response to the continuing crisis. The Federal Reserve further eased the terms of its securities lending program by reducing the minimum loan fee from 50 basis points to 10 basis points and by expanding the limit on total borrowings by a single dealer to $5 billion (Table 1). Treasury officials took the unprecedented step of reopening *four* off-the-run Treasury notes, announcing at 10:40 a.m. that they would "reopen multiple securities which have created severe dislocations in the market causing acute, protracted shortages."[53] Two of the reopened notes were auctioned later the same day (at 11:30 a.m. and 1:00 p.m., respectively), and the other two notes were auctioned the following day (Table 2). The decision to reopen a substantial amount ($10 billion each) of so many different notes made clear the scale of the fails problem; the decision to auction one note with less than an hour of notice and a second note with less than three hours of notice emphasized the urgency of the situation.

Although the reopenings helped to mitigate settlement fails in the issues that were reopened, aggregate primary dealer Treasury fails continued to rise, reaching a daily average level of $379 billion per day over the interval from October 9 to October 22 (Chart 3). Comments to the effect that "Treasury market functioning remains impaired" and "the repo market is not functioning" became commonplace. On October 17, a widely read market letter remarked that "the breakdown in the clearing mechanism for the Treasury market is beginning to emerge as a top-tier policy concern."[54]

[52] "Demand for Short-Term Treasury Debt Puts a Crimp in World-Wide Supply," *Wall Street Journal*, September 25, 2008, p. C1 (reporting that "some foreign central-bank officials … are reluctant to lend out their safest collateral—U.S. Treasurys."); "U.S. Treasury Steps Up Debt Sales to Reduce Shortages (Update 2)," Bloomberg.com, October 8, 2008, 12:43 EDT (quoting the head of interest rate strategy at Credit Suisse Securities as saying that "people are so nervous about the financial crisis that they're holding on to their collateral and not lending it out."); and "More Treasury Bonds on Way to Ease Crisis," *Wall Street Journal*, October 9, 2008, p. A6 (reporting that "investors have been unwilling to lend Treasury securities to other market participants.").

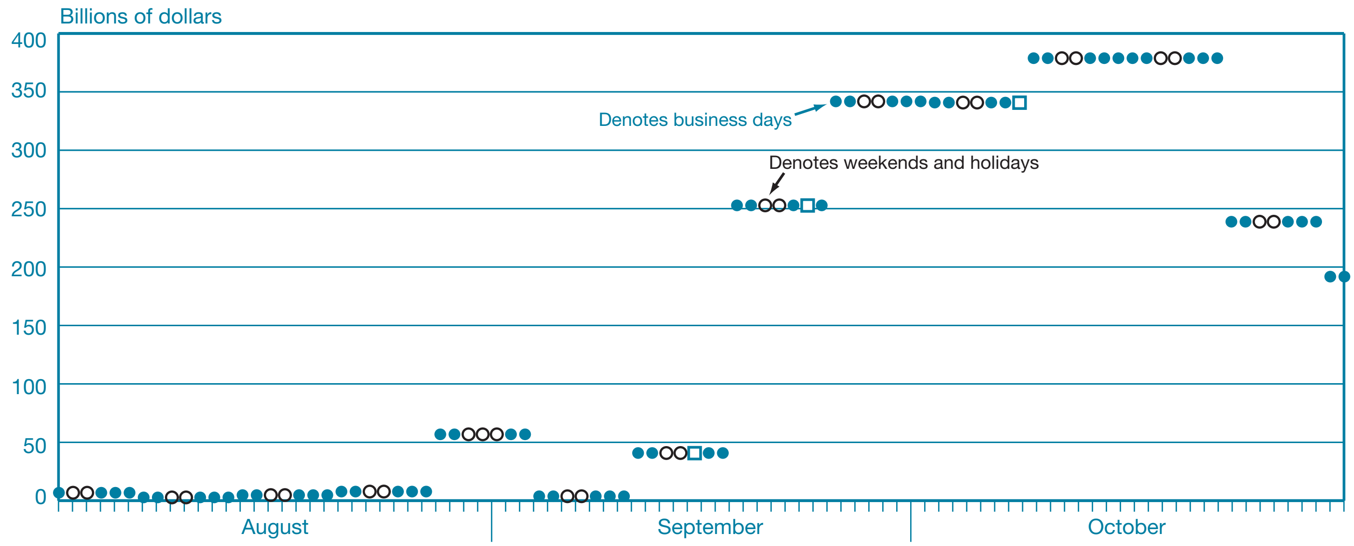
[53] U.S. Department of the Treasury, "Statement on Treasury Market Conditions and Debt Management Actions," October 8, 2008, available at http://www.treas.gov/press/releases//hp1186.htm.

[54] Wrightson, *Federal Reserve Data*, October 17, 2008.

CHART 3

Daily Average (Over Weekly Intervals) Primary Dealer Settlement Fails in Treasury Securities

August to October 2008



Source: Federal Reserve Bank of New York.

Notes: The first square marks the Lehman insolvency on Monday, September 15. The second square marks the effective date of a revision in the terms and conditions of the Federal Reserve System Open Market Account (SOMA) securities lending program on Tuesday, September 23. The third square marks the announcement of the surprise reopening of four Treasury notes on Wednesday, October 8, and the effective date of a further revision in the terms and conditions of the SOMA securities lending program on the same day.

TABLE 1

Terms and Conditions of Federal Reserve Security Loan Auctions

Effective Date	Theoretical Amount of a Single Issue Offered (Percentage of SOMA Holdings)[a]	Minimum Loan Fee (Basis Points)	Limits on Outstanding Borrowings by a Single Dealer: Per Issue	Limits on Outstanding Borrowings by a Single Dealer: Total
Terms and conditions prior to Lehman insolvency				
November 26, 2007	90	50	Lesser of $750 million and 25 percent of amount beneficially owned in SOMA portfolio	$3 billion
Post-Lehman terms and conditions				
September 23, 2008	90	50	Lesser of $750 million and 25 percent of amount beneficially owned in SOMA portfolio	**$4 billion**
October 8, 2008	90	**10**	Lesser of $750 million and 25 percent of amount beneficially owned in SOMA portfolio	**$5 billion**
December 18, 2008[b]	90	**1**	Lesser of $750 million and 25 percent of amount beneficially owned in SOMA portfolio	$5 billion

Source: Federal Reserve Bank of New York.

Notes: SOMA is the Federal Reserve System Open Market Account. Entries in bold indicate a change in terms.

[a]Amount actually offered is the lesser of the theoretical amount offered and the amount of the issue actually in the SOMA account at the time of an auction.

[b]Last revision prior to the end of 2008.

Table 2

Treasury Notes Reopened in October 2008

	4 1/8 Percent Note Maturing May 15, 2015	4 1/4 Percent Note Maturing August 15, 2015	4 Percent Note Maturing February 15, 2015	3 1/2 Percent Note Maturing February 15, 2018
Amount offered	$10 billion	$10 billion	$10 billion	$10 billion
Auction date and time	October 8, 2008, 11:30 a.m.	October 8, 2008, 1:00 p.m.	October 9, 2008, 11:30 a.m.	October 9, 2008, 11:30 a.m.
Issue date	October 15, 2008	October 15, 2008	October 15, 2008	October 15, 2008
Amount bid competitively	$12.1 billion	$21.1 billion	$23.7 billion	$23.1 billion
Closing market yield on October 7, 2008 (percent)	2.87	2.98	2.79	3.57
Auction yield (percent)	3.31	3.44	3.23	3.79
Closing market yield on October 9, 2008 (percent)	3.35	3.57	3.22	3.92

Sources: U.S. Treasury Department; *Wall Street Journal*.

Note: Over the interval from October 7 to October 9, the closing market yield on the on-the-run five-year note (the 3 1/8 percent note of September 30, 2013) rose from 2.47 percent to 2.79 percent and the closing market yield on the on-the-run ten-year note (the 4 percent note of August 2018) rose from 3.50 percent to 3.81 percent.

7. The TMPG Steps Up

By mid-October 2008, Treasury and Federal Reserve officials and private sector market participants understood that the volume and persistence of settlement fails in Treasury securities was a major problem, but what could or should be done was far from obvious. The four reopenings had reduced fails in the reopened notes, but speculation over whether the Treasury would reopen other chronically failing issues was contributing to unwanted volatility in the prices of other Treasury securities. Additionally, there was some indication that the reopenings had not been well received. The first auction, of $10 billion of the 4 1/8 percent notes of May 2015, attracted only $12.1 billion of tenders, and the notes were sold at a price almost 3 points below where outstanding notes of the same series traded prior to the auction.

An alternative approach was to revise the market convention of postponing—without any explicit penalty and at an unchanged invoice price—a seller's obligation to deliver Treasury securities if the seller failed to deliver the securities on a scheduled settlement date. However, precisely because the treatment of settlement fails was a matter of market convention, rather than law or regulation, it could not be changed except through widespread adoption of an alternative convention.

Fortuitously, in early 2007 the Federal Reserve Bank of New York had sponsored the organization of a new forum—the Treasury Market Practices Group—for discussing Treasury market practices and for advocating the adoption of practices deemed to be in the best interests of the market.[55]

The TMPG is a group of private sector market professionals committed to supporting the integrity and efficiency of the market for U.S. Treasury securities. Membership includes senior business managers and legal and compliance professionals from broker-dealer firms, banks, buy-side firms, and other organizations involved in Treasury market infrastructure. (Box 4 identifies the membership in October 2008.) The TMPG routinely meets about eight to ten times a year to discuss trading issues and best-practice recommendations for the Treasury market and publishes "Treasury Market Best Practices,"[56] a "living document" that aims to support

The TMPG is a group of private sector market professionals committed to supporting the integrity and efficiency of the market for U.S. Treasury securities.

[55] Federal Reserve Bank of New York, "Statement on Formation of Private-Sector Treasury Market Best Practices Group," February 9, 2007, available at http://www.newyorkfed.org/newsevents/news/markets/2007/an070209.html. See also Federal Reserve Bank of New York, "Statement Regarding New York Fed Meeting with Primary Dealers," November 6, 2006, available at http://www.newyorkfed.org/newsevetns/news-archive/markets/2006/an061105.html.

[56] Available at http://www.newyorkfed.org/tmpg/TMPG-best-practices_033109.pdf.

Box 4

Membership of the Treasury Market Practices Group in October 2008

Thomas Wipf, *Chair*	Morgan Stanley
Fran Bermanzohn	Goldman Sachs (last meeting in October 2008)
Arthur Certosimo	The Bank of New York Mellon
Daniel Dufresne	Citadel Investment Group, LLC
Peter Economou	State Street
John Fath	BTG
Michael Haddad	Caxton Associates (last meeting in January 2009)
Curt Hollingsworth	Fidelity Investments
James Hraska	Barclays Capital
Murray Pozmanter	Depository Trust & Clearing Corporation
Gerald Pucci	BlackRock
John Roberts	Barclays Capital
Bill Santangelo	Countrywide Securities Corp. (last meeting in October 2008)
Peter Stebbing	Reserve Bank of Australia
Nancy Sullivan	The Bank of New York Mellon
Matthew Zames	JPMorgan Chase

Subsequent Additions prior to May 1, 2009

Glenn Hadden	Goldman Sachs
Stuart Wexler	Merrill Lynch

Treasury market integrity and efficiency. Best-practice recommendations include guidelines for promoting market liquidity, for integrating compliance and trading functions in a meaningful fashion, and for managing large positions in ways that avoid adverse consequences for market liquidity. TMPG practice guidance has also addressed the efficient clearing and settlement of trades. Thus, the TMPG was well positioned in October 2008 to provide the leadership required to revise the market convention for settlement fails.[57]

The first meeting of the TMPG after the reopening auctions of October 8 and 9 was on Thursday, October 23. The chairman, Tom Wipf of Morgan Stanley, opened the meeting by reminding members of the urgency of the situation:

> To overstate the obvious, the work of today's meeting around settlement fails in Treasuries finds our committee at a crossroad. . . . At this critical juncture it is incumbent that TMPG take the leadership position on this issue and work as a group to provide practical, real time solutions. . . . Our goal as members of this committee is to support the integrity and efficiency of the U.S. Government Treasury Market. . . .

William Dudley, Executive Vice President of the Federal Reserve Bank of New York and Manager of the Fed's System Open Market Account, echoed Wipf's call for leadership: "This [meeting] is happening at a critical time in the market place where leadership is important to creating confidence and stability—we believe this group can, should and will provide that leadership."

7.1 The November 12 Recommendations

During the October 23 meeting, and in a series of subsequent telephone conference calls, TMPG members discussed changes in market practices that might reduce chronic fails and limit the likelihood of a recurrence. The group unveiled its recommendations on Wednesday, November 12, 2008.[58]

The principal recommendation suggested that "market participants agree that the invoice price . . . on any cash or financing transaction that fails to settle on the originally scheduled date be reduced at a fails [charge] rate equal to the greater of a) 3 percent per annum minus the fed funds target rate … and b) zero." As shown in Chart 4, this would penalize fails at a rate that starts at zero when the target federal funds rate is at or above 3 percent and rises to 3 percent as the target funds rate declines toward zero. It follows that the economic cost of failing would never fall below about 3 percent per annum.[59] The TMPG concluded that the "out-of-pocket cost to the party failing to deliver securities will provide a compelling incentive to resolve fails promptly."

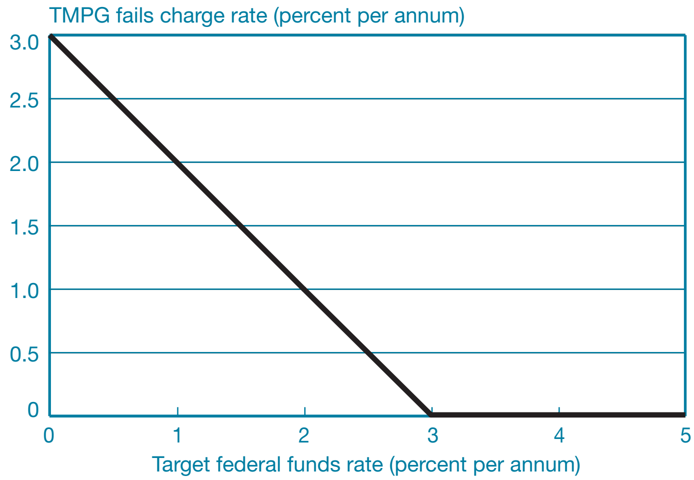
[57] The Association of Primary Dealers provided similar leadership in revising the market convention for the treatment of accrued interest in repurchase agreements after the 1982 failure of Drysdale Government Securities (Garbade 2006).

[58] Treasury Market Practices Group, "Treasury Market Practices Group Endorses Several Measures to Address Widespread Settlement Fails," November 12, 2008, available at http://www.newyorkfed.org/tmpg/PR081112.pdf. The recommendations were reported in "Treasury's Warning Adds to Plunge in Failed Trades," Bloomberg.com, November 12, 2008, 15:59 EST, and "Repo Experts Propose Plans to Counteract Rise in 'Fails,'" FT.com, November 12, 2008, 20:00.

[59] This follows because the sum of the target federal funds rate, which is usually at or slightly above the Treasury general collateral repo rate, and the fails charge rate would never be less than 3 percent. The TMPG could have referenced the overnight Treasury general collateral repo rate in lieu of the target federal funds rate, but the target funds rate is more familiar and more readily observable to market participants. There is no definitive and widely disseminated measure of overnight Treasury general collateral repo rates. The Federal Reserve, by comparison, publicly announces the target funds rate.

Chart 4

TMPG Fails Charge Rate as a Function of Target Federal Funds Rate



The TMPG explicitly based its recommendation on the dysfunctionality of the existing market convention for settlement fails:

> Past experience—for example, during the summer of 2003—shows that settlement fails in a particular [security] may become widespread and persistent when the special collateral repo rate for that [security] nears zero. Special collateral repo rates cannot exceed the Treasury general collateral repo rate. As a result, settlement fails across a wide variety of [securities] can . . . become widespread and persistent when the Treasury general collateral repo rate is near zero—as is currently the case.
>
> The underlying problem is the Treasury market contracting convention that a seller can deliver securities after the originally scheduled settlement date at an unchanged invoice price [and] without incurring any penalty. Introduction of a dynamic fails [charge] with a finite cap rate would remedy this problem. In particular, a dynamic fails [charge] would provide an incentive for sellers to resolve fails promptly, and could lead to repo contracting conventions [that is, negative repo rates] that would give beneficial owners of Treasury securities an opportunity to earn as much as the [3 percent] cap rate in securities loan fee income regardless of the level of nominal interest rates.[60]

[60] Treasury Market Practices Group, "Treasury Market Practices Group Endorses Several Measures to Address Widespread Settlement Fails," November 12, 2008.

The TMPG recognized that the "the introduction of [the recommended convention] raises operational, legal and other implementation issues that may vary across Treasury market participants" and promised to engage in "further analysis of these issues," with a goal of announcing by January 5, 2009, its recommendations for implementation.[61]

8. The Crisis Recedes but Support for Revising the Market Convention Persists

By the time the TMPG made its November 12 recommendation, settlement fails in the Treasury market were receding rapidly. As shown in Chart 5, primary dealer fails declined from a daily average of $379 billion during the week of October 16-22 to a daily average of $70 billion during the week of November 13-19 and averaged less than $50 billion a day in December.

Support for a revised market convention for settlement fails persisted in spite of the receding volume of fails, largely because the crisis of late September and early October had given added currency to the view that the existing convention was dysfunctional. The discussion of settlement fails during the November 4, 2008, meeting of the Treasury Borrowing Advisory Committee, as well as the views expressed in a prominent market newsletter in early January 2009, illustrates the growing consensus.[62]

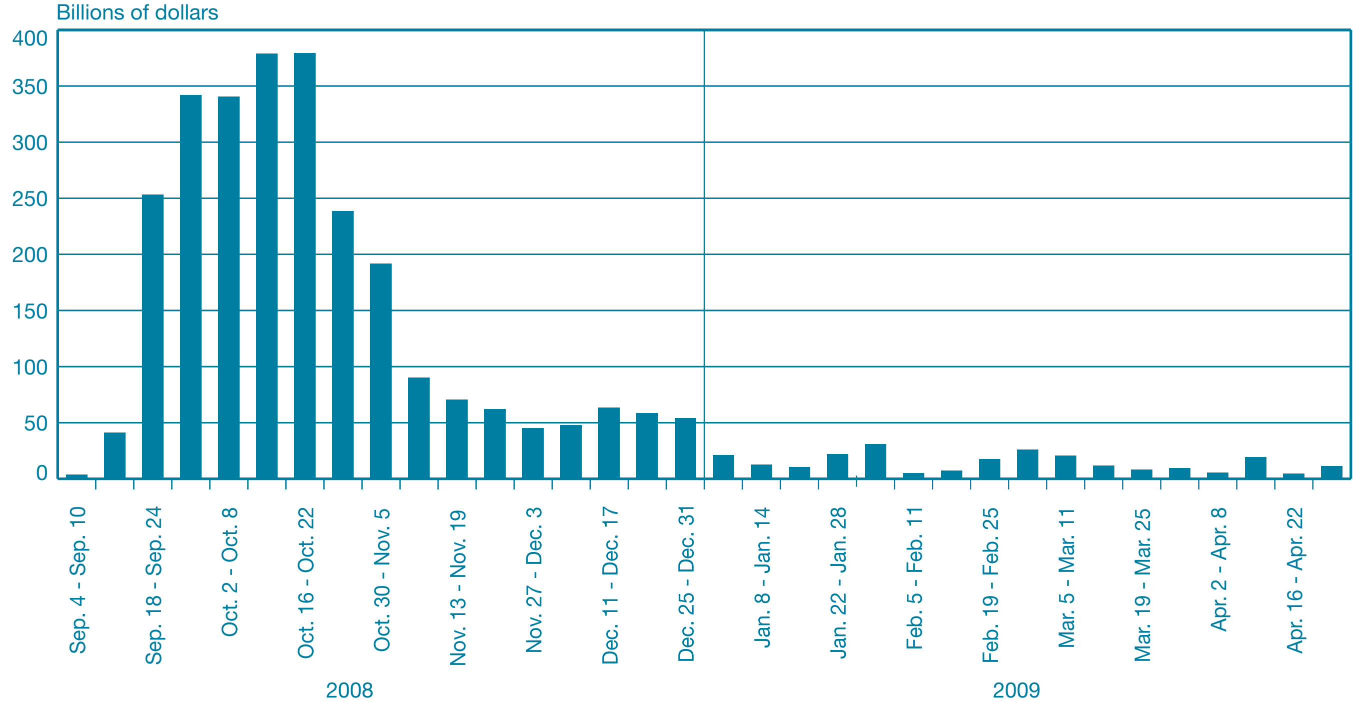
[61] The TMPG made three additional recommendations on November 12: 1) that market participants undertake a study of the most efficient way to margin fails in Treasury securities (in order to reduce counterparty credit risk exposure), 2) that market participants examine whether the Fixed Income Clearing Corporation, the two clearing banks, or other interested parties might develop "new or enhanced … multilateral netting arrangements" that might reduce settlement fails, and 3) that market participants pursue consensual cash settlement of transactions in Treasury securities that have been failing for more than five days. The TMPG also expressed its support for "discussion of a standing facility by the U.S. Department of the Treasury to provide temporary new supply of specific securities at a penalty rate when settlement fails persist," but noted that the creation of such a facility was a long-term goal and that progress on a fails charge should not be contingent on the development of a Treasury security lending facility.

[62] William Dudley, Executive Vice President, Federal Reserve Bank of New York, and Manager of the System Open Market Account, stated during a public conference call on January 14, 2009, on the TMPG fails initiative that: "Although settlement fails have declined recently from record levels amid reduced trading volumes, the extremely low level of interest rates suggests that fails could again rise significantly when trading activity picks up. The fundamental incentive to deliver securities under current market conditions is simply not sufficient at very low nominal interest rates to reduce the probability of large chronic fails to acceptable levels."

Chart 5

Daily Average (Over Weekly Intervals) Primary Dealer Settlement Fails in Treasury Securities

September 2008 to April 2009



Source: Federal Reserve Bank of New York.

8.1 The November Meeting of the Treasury Borrowing Advisory Committee

The Treasury Borrowing Advisory Committee (TBAC) is a committee of market professionals selected to advise the Secretary of the Treasury on matters relating to Treasury debt management. At its November 4 meeting, the committee discussed the upcoming midquarter refunding and, *inter alia*, the fails situation.

The TBAC's discussion of settlement fails focused initially on better ways for the Treasury to reopen outstanding issues than the "snap" reopenings of October 8 and 9, but then turned to the market convention for settlement fails. Several committee members observed that investors had "little economic incentive to lend securities when general collateral [repo] rates stood at 20 basis points," and one member suggested that "a negative [repo] rate of 200 or 300 basis points . . . would create the correct economic incentives to cause holders of securities in low interest rate environments to lend securities again."[63]

In its ensuing report to the Secretary of the Treasury, the TBAC expressed the view that the low level of short-term interest rates "has made the cost of failing negligible, [leaving] little desire for short-sellers to close out their positions" and noted the suggestion of one committee member "that there should be a cost in the form of a penalty rate associated with fails in a low-rate environment." The report further noted that such a cost would encourage negative-rate repo trading, "which would allow the free market to determine the effective cost of the fail, and change the economics of securities lending."[64]

[63] U.S. Department of the Treasury, "Minutes of the Meeting of the Treasury Borrowing Advisory Committee of the Securities Industry and Financial Markets Association, November 4, 2008," November 5, 2008, available at http://www.ustreas.gov/press/releases/hp1239.htm.

[64] U.S. Department of the Treasury, "Report to the Secretary of the Treasury from the Treasury Borrowing Advisory Committee of the Securities Industry and Financial Markets Association, November 4, 2008," November 5, 2008, available at http://www.ustreas.gov/press/releases/hp1238.htm.

8.2 Comments in Wrightson's *Money Market Observer*

The January 5, 2009, edition of Wrightson's *Money Market Observer* devoted substantial space to the TMPG proposal to revise the market convention for settlement fails. The newsletter noted that creating an explicit charge for settlement fails was "the most straight-forward way to remedy the obvious structural flaws that lead to delivery logjams in today's market," and pointed out the anticipated benefits of restoring competitive market forces to the special collateral repo markets: "The TMPG believes (correctly, in our view) that the repo market will be more elastic—and the Treasury clearing process more efficient—if the floor on special repo rates is set low enough [that is, below zero] to preserve a spread relative to general collateral rates even in the current rate environment."

9. Getting It Right

Although a consensus had emerged in support of revising the market convention for settlement fails, the TMPG recommendation for reducing invoice prices itself needed some revision.

The TMPG recommendation would have required a seller and a buyer to reduce the invoice price on a failing transaction by *matching amounts* on a *daily* basis. If one party reduced the invoice price and the other did not, or if the two parties reduced the invoice price by different amounts, any attempt by the seller to deliver securities against payment would be rejected by the buyer (because the buyer would be looking to pay a different amount than what the seller was looking to receive). TMPG members who understood the complex architecture of broker-dealer and custodian settlement systems pointed out that requiring matching daily price reductions would impose a major operational burden on market participants and could lead to an explosion in rejected deliveries (and thus in settlement fails).

Although a consensus had emerged in support of revising the market convention for settlement fails, the TMPG recommendation for reducing invoice prices itself needed some revision.

In lieu of adjusting invoice prices, several TMPG members suggested that an economically equivalent result could be obtained if a seller who makes a late delivery agrees to make a *side payment* to the buyer in an amount equal to what became known as the "TMPG fails charge." The charge for a fail on a given business day would be computed as:

$$(1) \qquad C = \frac{n}{360} \times .01 \times max[3 - R_{trgt}, 0] \times P,$$

where:

C = charge, in dollars,

n = number of calendar days to the next following business day,

R_{trgt} = target federal funds rate at the close of business on the business day preceding the fail, in percent per annum, and

P = total proceeds due from the buyer, in dollars.

For example, if P = \$10 million, R_{trgt} = 1 percent, and n = three days, then C = \$1,666.66.[65] This procedure had the advantage of not requiring any change in existing settlement systems.

The idea of replacing invoice price adjustments with side payments illustrates an important aspect of the TMPG initiative: by working collaboratively, the TMPG was able to achieve its objectives while accommodating an existing institutional structure: back-office settlement systems. The difference between a price adjustment and a side payment may seem trivial, but the success of the TMPG initiative hinged on recognizing the difference.

9.1 The January 5 Announcement

On January 5, 2009, the TMPG announced that it was recommending a fails charge in the form of a side payment on transactions that failed to settle on a timely basis and that it was making several additional refinements to its November 12 recommendation.[66] The three key refinements

[65] \$1,666.66 = (3/360) × .01 × *max*[3 – 1, 0] × \$10,000,000.

[66] Treasury Market Practices Group, "Timeline for New Market Practices to Address Widespread Settlement Fails in U.S. Treasury Securities," January 5, 2009, available at http://www.newyorkfed.org/tmpg/PR090105c.pdf, and Treasury Market Practices Group, "Claiming a Fails Charge for a Settlement Fail in U.S. Treasury Securities," January 5, 2009, available at http://www.newyorkfed.org/tmpg/PR090105a.pdf. The January 5 announcement was reported in "Repo Arena Gets a Plan on Penalties," *Wall Street Journal*, January 6, 2009, p. C3, and "Penalty for Failed Trades Set to Spark New Era for US Repo," *Financial Times*, January 7, 2009, p. 22.

were 1) a statement of the process for claiming a fails charge, 2) a timeline suggesting that market participants begin claiming for settlement fails on transactions agreed to on or after May 1, 2009, and 3) replacement of the target federal funds rate (in the formula for the fails charge, equation 1 above) with a "TMPG reference rate." The latter rate was defined as the target federal funds rate if the Federal Open Market Committee specified a target rate or the lower limit of the target band for the federal funds rate if the FOMC specified a target band.[67] In the event the FOMC specified neither a target rate nor a target band, the TMPG committed to recommending some other similar, readily observable, short-term interest rate as a reference rate for the fails charge formula.[68]

On January 5, 2009, the TMPG announced that it was recommending a fails charge in the form of a side payment on transactions that failed to settle on a timely basis and that it was making several additional refinements to its November 12 recommendation.

The decision to recommend a side payment (in lieu of an invoice price adjustment) required the TMPG to specify a way for buyers to collect from sellers who failed to deliver securities on a timely basis. In the case of buyers and sellers who settled through FICC, a collection process could be added to other similar processes previously implemented by FICC (such as the collections and disbursements that result from marking transactions to current market prices).[69] However, the collection process was not as simple for transactions that settled bilaterally, as was the case for most transactions between dealers and their nondealer customers.

[67] This charge was necessitated by the December 16, 2008, decision of the Federal Open Market Committee to establish a target range for the federal funds rate of 0 to 1/4 percent.

[68] In late March 2009, the TMPG announced a slightly different form for the fails charge computation:

$$C = \frac{1}{360} \times .01 \times max[3 - R_{TMPG}, 0] \times P .$$

In this form, the charge is computed for each *calendar* day that a seller's delivery obligation is failing. R_{TMPG} is the TMPG reference rate on the business day preceding the day for which the charge is computed. See Treasury Market Practices Group, "Treasury Market Practices Group Announces Updates to Fails Charge Recommendation," March 30, 2009, available at http://www.newyorkfed.org/tmpg/tmpg_033009.pdf.

The January 5 announcement suggested that the best way to initiate the fails charge would be for buyers to tender claims directly to sellers.[70] A seller could either pay what was claimed or dispute the claim and negotiate with its counterparty over the amount due.

The TMPG further suggested that if an investor employed a professional asset manager and that manager contracted to sell securities that were not delivered on a timely basis, the claim for the fails charge should be directed to the asset manager (rather than to the investor or to the investor's custodian). This suggestion was based on the pragmatic notion that since the sale had been negotiated by the asset manager, the asset manager would be in the best position to recognize the sale and identify who was responsible for the settlement fail, be it the asset manager, the investor's custodian, or some other party,[71] or whether the claim should be left for the account of the investor.[72]

9.2 Trading Practice and Market Practice Recommendations

Following the January 5 announcement, TMPG members and other market participants collaborated to publish two documents providing guidance on how to implement the TMPG fails charge. The documents were important for clarifying how fails charges should be calculated and claimed and generally for enhancing the transparency of the new market convention.

[69] The addition required a change in FICC rules that had to be approved by the Securities and Exchange Commission. FICC filed the proposed rule change on February 25, 2009 (Securities and Exchange Commission Release no. 34-59569, March 12, 2009), and the Commission granted approval two months later (Securities and Exchange Commission Release no. 34-59802, April 20, 2009).

[70] The January 5 announcement noted the possibility of setting up a central industry utility to receive and process claims, but observed that the design of such a facility raised novel questions regarding the identification of buyers and sellers and would require further consultation with market participants.

[71] Some investors retain an agent to lend securities from the investor's portfolio. Such agents are commonly called "agent sec lenders." In most cases, an agent sec lender is obliged to reclaim securities out on loan if the investor's asset manager decides to sell the securities. If an agent sec lender fails to reclaim securities on a timely basis and thereby causes a settlement fail, the fail may be the responsibility of the agent sec lender, rather than the asset manager or the custodian.

[72] A sale of securities negotiated by an asset manager may fail to settle on a timely basis because the investor's custodian failed to receive the same securities on an unrelated purchase. Such fails cannot be attributed to faulty behavior by the asset manager or the investor's custodian, so the resulting fails charge would be left for the account of the investor. The investor can, of course, direct its asset manager to file a claim on the seller who failed to deliver securities to the investor.

Trading practice recommendations: On January 15, 2009, the TMPG and the Securities Industry and Financial Markets Association (SIFMA) published a "U.S. Treasury Securities Fails Charge Trading Practice"[73] to give market participants guidance on exactly the types of transactions that were covered by, and excluded from, the TMPG fails charge. The "Trading Practice" also recommended the form of a letter that a market participant could send to counterparties, advising them of the participant's adoption of the new policy for settlement fails, and suggested a statement that could be added to trade confirmations indicating that a transaction was subject to the fails charge.

TMPG members and other market participants collaborated to publish two documents providing guidance on how to implement the TMPG fails charge. The documents were important for clarifying how fails charges should be calculated and claimed and generally for enhancing the transparency of the new market convention.

Market practice recommendations: On April 23, 2009, SIFMA published a "Treasury Market Practices Group Fails Charge Market Practice"[74] that recommended procedures for buy-side firms to use in connection with the new fails charge. The recommended procedures included processes for researching and tracking fails, calculating fails charges, determining responsibility for a claim, sending and receiving claims, and accounting for claims. The procedures also included suggestions made earlier by SIFMA and adopted by the TMPG[75] that claims be submitted at the beginning of a month for fails settled during the preceding month (to accommodate custodians and asset managers who structured their control systems around settled transactions) and be in excess of $500 per issue per settlement (to limit costly research and billing efforts to nontrivial claims).

10. Implementation

The TMPG fails charge went into effect on May 1, 2009, replacing the former market convention of postponing—without any explicit penalty and at an unchanged invoice price—a seller's obligation to deliver Treasury securities when the seller fails to deliver the securities on a scheduled settlement date. Henceforth, the cost of failing to settle a sale of Treasury securities in a timely fashion would not be less than 3 percent per annum.

It would be premature to argue that the TMPG fails charge has eliminated the possibility of yet another episode of chronic settlement fails in Treasury securities; past episodes were rare to begin with and some future event may demonstrate the existence of an unsuspected flaw in the new system. It may also be the case that the 3 percent benchmark rate is too low and that chronic fails would be better mitigated with a 3 1/2 or 4 percent rate. However, there is no evidence to date that the new market convention, and the 3 percent benchmark rate, are not working. Chart 6 shows daily average settlement fails over weekly intervals from the beginning of 2009 to July 2010. Fails averaged a bit over $14.4 billion per day during the first four months of 2009, but only $4.2 billion per day since implementation of the fails charge.[76] More important, the relatively modest eruptions of settlement fails that appeared during the first week of July 2009 and the first week of January 2010 quickly subsided. The new convention is not yet out of its infancy, but there is reason to anticipate that the TMPG fails charge will similarly dampen future eruptions.

The new convention is not yet out of its infancy, but there is reason to anticipate that the TMPG fails charge will . . . dampen future eruptions [of settlement fails].

[73] Treasury Market Practices Group and Securities Industry and Financial Markets Association, "U.S. Treasury Securities Fails Charge Trading Practice," January 15, 2009, available at http://www.sifma.org/capital_markets/docs/Fails-Charge-Trading-Practice.pdf.

[74] Securities Industry and Financial Markets Association, "Treasury Market Practices Group Fails Charge Market Practice," April 23, 2009, available at http://www.theasset manager.com/docs/AM_Custodian_IndustryProcedures_TMPG_FailsCharge.pdf.

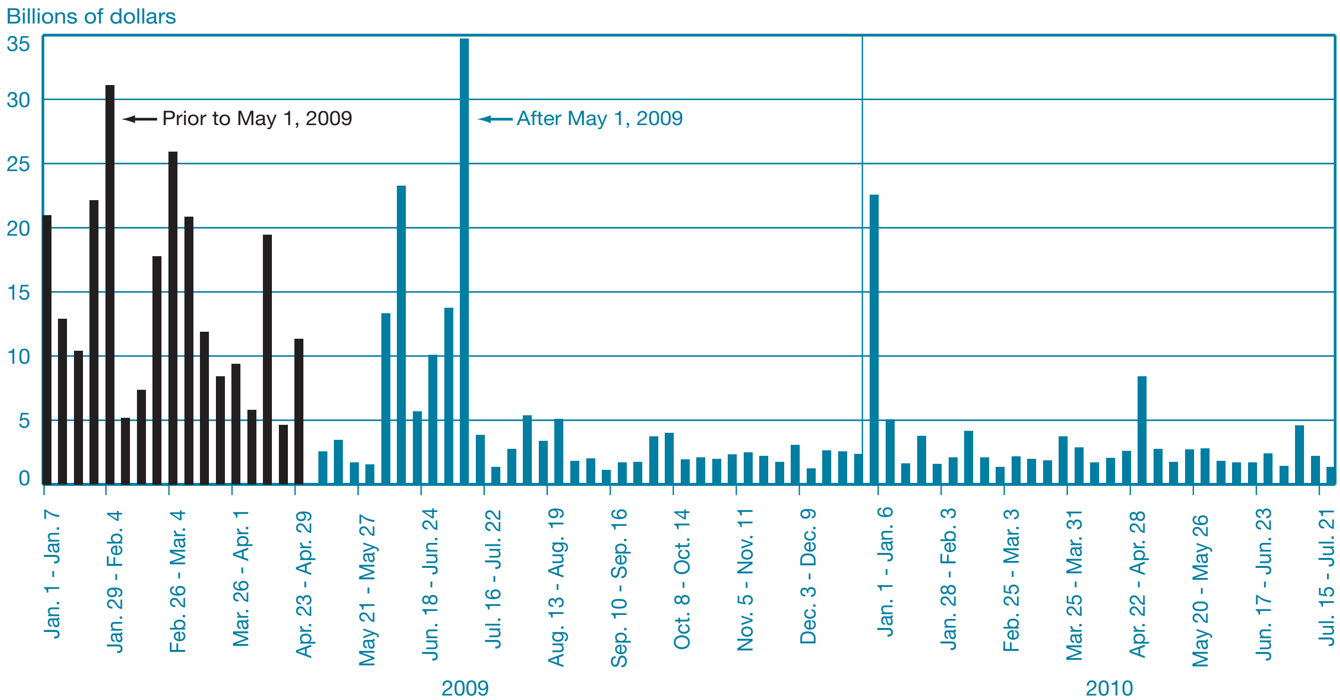
[75] See Treasury Market Practices Group, "Treasury Market Practices Group Announces Updates to Fails Charge Recommendation," March 30, 2009, available at http://www.newyorkfed.org/tmpg/tmpg_033009.pdf.

[76] The fails charge was never intended to eliminate *all* settlement fails. (Fails attributable to miscommunication or operational problems are unlikely to be eliminated by the fails charge—although they may be resolved more quickly.) Rather, the fails charge was aimed at mitigating episodes of *chronic* fails that can threaten market liquidity and efficiency.

Chart 6

Daily Average (Over Weekly Intervals) Primary Dealer Settlement Fails in Treasury Securities

January 2009 to July 2010



Source: Federal Reserve Bank of New York.

11. Conclusion

The TMPG fails charge initiative is important both for what it accomplished and for how it was accomplished. Substantively, the initiative revised an outmoded convention and mitigated an important dysfunctionality in a market of critical national significance. Procedurally, the initiative demonstrated how cooperation between the public and private sectors can speed innovative and efficient responses to changing circumstances. At the time of the May 1, 2009, implementation of the fails charge, the Federal Reserve Bank of New York welcomed the new convention:

> We applaud the dedicated efforts of the TMPG in spearheading the development and implementation of this targeted solution to the settlement fails problem. This significant milestone in the evolution of Treasury market practice demonstrates that groups, such as the TMPG, are effective in addressing deficiencies in market functioning and facilitating market best practices.[77]

In a subsequent letter to the TMPG membership expressing his personal thanks for the Group's dedication and commitment to making the fails charge a reality, William Dudley, now president of the Federal Reserve Bank of New York, reflected on the significance of the new market convention:

> The implementation of the fails charge marks a rare and significant evolution in Treasury market architecture. In my view, one would need to look back to 1982 to find a development of similar magnitude, when the collapse of Drysdale Securities led to the adoption of a new market practice to include accrued interest in repo contracts. The fails charge stands among relatively few revisions to contracting conventions in the Treasury market since the development of a liquid national market following World War I.

[77] Federal Reserve Bank of New York, "New York Fed Applauds Implementation of the TMPG's Fails Charge Recommendation," May 1, 2009, available at http://www.newyorkfed.org/newsevents/news/markets/2009/ma090501.html.

Appendix: Additional Individuals Who Provided Feedback and Assistance in the Implementation Phase of the TMPG Fails Charge Initiative

The following list attempts to include all non-TMPG (Treasury Market Practices Group) individuals who either served on formal subgroups that contributed to the timely implementation of the TMPG fails charge or participated in less formal conference calls and meetings. The authors apologize for any inadvertent omissions.

David Aman	Cleary Gottlieb Steen & Hamilton LLP
Marc Baum	Ramius LLC
Brandon Becker	WilmerHale
Brent Blake	State Street
Gary Buki	The Bank of New York Mellon
Kevin Caffrey	The Bank of New York Mellon
Maria Carina	Euroclear
Michael Cetta	AllianceBernstein Holdings
Brayton Cherry	Brown Brothers Harriman & Co.
Ed Corral	JPMorgan Chase
David Cosgrove	ICAP
Brian Crowe	Fidelity Investments
Craig Delany	JPMorgan Chase
Laura Dietel	State Street
Craig Dudsak	Citigroup
Steve Dunn	JPMorgan Chase
Joe Finan	Morgan Stanley & Co.
Marcellus Fisher	PIMCO
Barbara Friedman	New York Life Investment Management LLC
Robert Good	Goldman, Sachs & Co.
Simon Griffiths	JPMorgan Chase
Olivier Grimonpont	Euroclear
Chris Haas	Merrill Lynch
Janice Hamilton	The Northern Trust Company
Robert Hennessy	The Bank of New York Mellon
Eugene Ing	Depository Trust & Clearing Corporation
Dyann Kiessling	Fidelity Investments
Bradley Koehler	Euroclear
Marty Kruse	BNY Mellon Asset Servicing
Michael Landolfi	State Street
Lourdes Leon	Morgan Stanley & Co.
Christine Lin	Citigroup
Fred Lipinski	Depository Trust & Clearing Corporation
Colin Lloyd	Cleary Gottlieb Steen & Hamilton LLP
Frank Lupica	Lord, Abbett & Co. LLC
Diane Madera	Morgan Stanley & Co.
Claudia Maia	Euroclear
Jennifer Manor	Fidelity Investments
Frank Martino	New York Life Investment Management LLC
Christopher Marzullo	Lord, Abbett & Co. LLC
Jason McCann	Reserve Bank of Australia
Shirley McCoy	JPMorgan Chase
Katherine McGaugh	State Street
Kevin Meagher	Fidelity Investments
Omar Medina	Goldman Sachs Asset Management

Appendix: Additional Individuals Who Provided Feedback and Assistance in the Implementation Phase of the TMPG Fails Charge Initiative (Continued)

Fatima Mehladi	Euroclear
Eric Miller	Citadel Solutions, LLC
Tamara Molinary	AllianceBernstein Holdings
Carolyn Monroe-Koatz	JPMorgan Chase
Penny Morgan	Western Asset Management
Louis Nazarro	JPMorgan Chase
Edward Neeck	JPMorgan Chase
Peter Novello	Lord, Abbett & Co. LLC
Elisa Nuottajarvi	SIFMA
Paul Parseghian	Prudential Investment Management Inc.
Judith Polzer	JPMorgan Chase
Joseph Pomo	Goldman Sachs Asset Management
Thomas Ponti	State Street
Nancy Prior	Fidelity Investments
Pawan Puneet	State Street IMS
Christopher Ramsay	Citadel Investment Group, LLC
Bill Rose	BTG
Theodore Rothschild	JPMorgan Chase
Timothy Ryan	SIFMA
Joseph Sack	SIFMA
Randy Snook	SIFMA
Guido Stroemer	UBS
Brian Swann	Goldman Sachs
Rob Toomey	SIFMA
Diane Trupia	SIFMA
Raymond Tyrrell	Brown Brothers Harriman & Co.
Jason Ward	Fidelity Investments
Andrew Waskow	Goldman Sachs
Mark Willis	Merrill Lynch
Patricia Yak	Credit Suisse First Boston
Lawrence Young	Credit Suisse First Boston
Anthony Zook	JPMorgan Chase

References

Committee on Banking, Finance, and Urban Affairs. 1985. "The Federal Reserve Bank of New York Discount Window Loan of $22.6 Billion Extended to the Bank of New York." Hearing before the Subcommittee on Domestic Monetary Policy of the Committee on Banking, Finance, and Urban Affairs. U.S. House of Representatives, 99th Cong., 1st Sess., December 12.

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

Department of the Treasury. 2006. "Consideration of a Proposed Treasury Securities Lending Facility." Federal Register 71, no. 85 (May 3): 26,174-8.

Department of the Treasury, Securities and Exchange Commission, and Board of Governors of the Federal Reserve System. 1992. "Joint Report on the Government Securities Market." January.

Duffie, D. 1996. "Special Repo Rates." Journal of Finance 51, no. 2 (June): 493-526.

Fisher, M. 2002. "Special Repo Rates: An Introduction." Federal Reserve Bank of Atlanta Economic Review 87, no. 2 (second quarter): 27-43.

Fleming, M., and K. Garbade. 2002. "When the Back Office Moved to the Front Burner: Settlement Fails in the Treasury Market after 9/11." Federal Reserve Bank of New York Economic Policy Review 8, no. 2 (November): 35-57.

———. 2003. "The Repurchase Agreement Refined: GCF Repo." Federal Reserve Bank of New York Current Issues in Economics and Finance 9, no. 6 (June).

———. 2004. "Repurchase Agreements with Negative Interest Rates." Federal Reserve Bank of New York Current Issues in Economics and Finance 10, no. 5 (April).

———. 2005. "Explaining Settlement Fails." Federal Reserve Bank of New York Current Issues in Economics and Finance 11, no. 9 (September).

———. 2007. "Dealer Behavior in the Specials Market for U.S. Treasury Securities." Journal of Financial Intermediation 16, no. 2 (April): 204-28.

Garbade, K. 2004. "Origins of the Federal Reserve Book-Entry System." Federal Reserve Bank of New York Economic Policy Review 10, no. 3 (December): 33-50.

———. 2006. "The Evolution of Repo Contracting Conventions in the 1980s." Federal Reserve Bank of New York Economic Policy Review 12, no. 1 (February): 25-40.

———. 2007. "The Emergence of 'Regular and Predictable' as a Treasury Debt Management Strategy." Federal Reserve Bank of New York Economic Policy Review 13, no. 1 (March): 53-71.

Garbade, K., and J. Ingber. 2005. "The Treasury Auction Process: Objectives, Structure, and Recent Adaptations." Federal Reserve Bank of New York Current Issues in Economics and Finance 11, no. 2 (February).

Garbade, K., and J. Kambhu. 2005. "Why Is the U.S. Treasury Contemplating Becoming a Lender of Last Resort for Treasury Securities?" Federal Reserve Bank of New York Staff Reports, no. 223, October.

Ingber, J. 2006. "A Decade of Repo Netting." Futures & Derivatives Law Report 26, no. 2 (February): 3-12.

Jordan, B., and S. Jordan. 1997. "Special Repo Rates: An Empirical Analysis." Journal of Finance 52, no. 5 (December): 2051-72.

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

Sender, H. 1986. "The Day the Computers Went Down." Institutional Investor 11, no. 3 (March): 203-7.