Center for STUDYING DISABILITY POLICY

Initial Impacts of the Ticket to Work Program for Young New Social Security Disability Awardees: Estimates Based on Randomly Assigned Mail Months

Final Report

July 30, 2013

David Stapleton
Arif Mamun
Jeremy Page

MATHEMATICA Policy Research

This page has been left blank for double-sided copying.

Center for STUDYING DISABILITY POLICY

Contract Number:
0600-03-60130

Mathematica Reference Number:
08977.963

Submitted to:
Social Security Administration
Office of Retirement and Disability Policy
500 E Street SW, 9th Floor
Washington, DC 20024
Project Officer: Paul O'Leary

Submitted by:
Mathematica Policy Research
1100 1st Street NE, 12th Floor
Washington, DC 20002-4221
Telephone: (202) 484-9220
Facsimile: (202) 863-1763
Project Director: Gina Livermore

Initial Impacts of the Ticket to Work Program for Young New Social Security Disability Awardees: Estimates Based on Randomly Assigned Mail Months

Final Report

July 30, 2013

David Stapleton
Arif Mamun
Jeremy Page

MATHEMATICA Policy Research

This page has been left blank for double-sided copying.

CONTENTS

EXHIBITS

This page has been left blank for double-sided copying.

ACRONYMS

AV	All Variation
CDR	Continuing Disability Review
DAC	Disabled Adult Children
DF	Degrees of Freedom
DWB	Disabled Widow(er) Beneficiary
EN	Employment Network
EPE	Extended Period of Eligibility
IME	Indexed Monthly Earnings
IMM	Intended Mail Month
ITT	Intent-to-Treat
IV	Instrumental Variables
MIE	Medical Improvement Expected
MM	Actual Mail Month
NBS	National Beneficiary Survey
NSTW	Nonpayment Status Following Suspension or Termination for Work
NY	New York
OASI	Old Age and Survivor Insurance
PIA	Primary Insurance Amount
SE	Standard Error
SGA	Substantial Gainful Activity
SRV	Service Enrollment
SSA	Social Security Administration
SSD	Social Security Disability (includes disabled worker beneficiaries, DAC and DWB)

SSDI	Social Security Disability Insurance
SSI	Supplemental Security Income (Title XVI of the Social Security Act)
SSN	Social Security Number
STW	Suspended or Terminated for Work
SVRA	State Vocational Rehabilitation Agency
TESG	Technical Expert Support Group
TOT	Treatment-on-the-Treated
TRF	Ticket Research File
TTW	Ticket to Work
TWP	Trial Work Period

EXECUTIVE SUMMARY

This report presents results from a new analysis of the impacts of introducing the original Ticket to Work (TTW) program, the program that was in place before the regulatory changes of July 2008. An earlier analysis produced evidence that TTW had positive impacts on service enrollment (that is, enrollment in services at a state vocational rehabilitation agency or other employment network), but methodological limitations led to ambiguous conclusions about impacts on earnings and benefits (Thornton et al. 2007; Stapleton et al. 2008). The new analysis substantially resolves these ambiguities.

Substantial Innovations in Data and Methods

The new analysis incorporates multiple innovations relative to the earlier analysis, including the following:

- **Exploiting variation in when Tickets were mailed within each of the three rollout phases.** The earlier analysis relied heavily on the fact that SSA implemented the program in phases to three groups of states, using beneficiaries in late rollout states as a contemporaneous comparison group for beneficiaries in early rollout states. That turned out to be problematic because of differential trends in outcomes across phase groups that predated TTW. The new analysis exploits a feature of each phase's rollout: just before the start of the rollout, SSA selected the month in which it intended to mail each eligible beneficiary's Ticket (hereafter the "intended mail month", IMM) in an essentially random fashion. We are able to use this random variation to rigorously estimate how the timing of ticket mailing affects outcomes over a 48-month period after the rollout start. Those estimates can be used to infer the impacts of mailing the ticket at the beginning of the rollout versus not mailing it at all.
- **Exploiting improvements in the measurement of work outcomes.** Whereas the earlier analysis accounted only for annual earnings outcomes based on Internal Revenue Service (IRS) data in SSA's Master Earnings File, the new analysis allowed us to examine recently developed monthly outcome variables for benefit suspension or termination for work (STW) and for months in nonpayment status after STW (NSTW months). We were also able to estimate impacts on monthly outcomes for service enrollment, start of the Trial Work Period (TWP), and TWP completion.
- **Focusing on a subgroup of beneficiaries who, based on earlier findings, seem the most likely to use their Ticket to support their work efforts:** new Social Security Disability (SSD) beneficiaries younger than 40 who are not also receiving Supplemental Security Income (that is, SSD-only beneficiaries).
- **Examining outcomes over a longer period.** The earlier analysis accounted for impacts in two years only—the rollout year and the year after rollout. This might not have allowed enough time for substantial impacts on earnings and benefits to emerge. With the passage of time and the availability of monthly variables, we could, through the new analysis, estimate impacts on outcomes over 48 months following the start of the rollout in each phase.

The new analysis directly addresses the following research questions. Each question concerns the impact of duration from the month before the rollout start in a beneficiary's state to the month when the Ticket was mailed on outcomes over the 48 months after the rollout start.

- Did an increase in duration from the rollout start to Ticket mailing reduce the extent to which each of four "event" outcomes (service enrollment, TWP start, TWP completion, and STW) were attained as of 12, 24, 36, and 48 months after rollout start?
- Did an increase in duration from the rollout start to Ticket mailing reduce the number of months accumulated in NSTW months as of 12, 24, 36, and 48 months after rollout start?

Of course, it is not possible to directly estimate a more interesting set of impacts, namely the impacts of mailing Tickets versus never mailing Tickets—that is, of rolling out the TTW program versus the counterfactual of never rolling it out. Direct estimation is not feasible because SSA eventually mailed a Ticket to essentially all eligible beneficiaries in each phase's sample. Thus, there is no traditional "control" group in each phase's sample; instead, SSA's intent was to offer the treatment (that is, mail the Ticket) to each group, but at different times.

Nonetheless, estimates of the impacts of duration to mail month on the outcomes have important implications for the impacts of mailing the Tickets versus never mailing them, and can even be used to infer what the latter impacts are. Evidence that duration to mail month has an impact on an outcome over this period is evidence that mailing the Ticket has an impact on that outcome relative to never mailing the Ticket. Further, under reasonable assumptions, we can infer the total impact of mailing the Ticket, versus never mailing it, by estimating the impact of duration to mail month on outcomes measured at multiple points in after the rollout start (12, 24, 36, and 48 months later).

The methodology makes adjustments for the fact that SSA did not mail all Tickets according to the schedule. Some were mailed earlier, under a policy called "Ticket-on-demand," and others were not mailed at all because the beneficiary became ineligible between the date on which mail months were selected and the mail month itself—in a majority of cases because the beneficiary had died. The number of cases in which Tickets were not mailed in the intended month is small as a share of each phase's sample, but simply ignoring this fact would bias the impact estimates of duration to mail month. Fortunately, there is a well-established statistical approach to addressing this issue; we use the IMM as an *instrumental variable* for the actual mail month (MM).

The remainder of this summary focuses on the findings for Phases 2 and 3 because the methodology's ability to detect impacts is much greater for these phases than for Phase 1. One reason is that the Phase 1 sample had to be split into two relatively small samples because an operational issue led to different rollout schedules for New York and the rest of the Phase 1 states. A second reason is that the rollouts Phases 2 and 3 (11 months in each) were substantially longer than in either part of Phase 1 (nine months in New York and five months in the rest of Phase 1).

Strong Impacts on Service Enrollment, No Consistent Evidence of Impacts on Other Outcomes

We found clear evidence that the Ticket mailings during the rollout period significantly increased service enrollment. However, we found little evidence that this impact translated into a substantive increase in STW or in the number of NSTW months. Key impact estimates of duration

to MM for the latter two outcomes from the Phase 2 rollout are marginally significant and consistent with a substantive impact. But findings from other phases and for intermediate outcomes (TWP start and completion) in Phase 2 fail to confirm the Phase 2 findings; in fact, other findings suggest that the Phase 2 findings are simply due to chance. Projections of total impacts at 48 months reinforce this interpretation.

The strongest findings pertain to how the duration from rollout start to Ticket mailing affects service enrollment in the first 12 months. The Phase 2 and 3 findings show not only that the impacts are significant and positive but also that they are consistent with each other. The Phase 2 point estimate implies that that mailing a ticket in the first rollout month versus waiting until month 13 results in a 1.0 percentage point increase in service enrollment as of the end of month 12. The Phase 3 estimate implies a 0.8 percentage points impact. These estimates are substantial. For both Phase 2 and 3, the estimated impacts are equivalent to 15 percent of the mean percentage of beneficiaries in the sample who enrolled in services over the entire 48-month observation period.

Thus, it is clear from these results that the intended delay in mailing the ticket to some beneficiaries delayed the impact on service enrollment. The results also indicate, however, that the delay did not reduce the eventual impact on service enrollment. This is evident from the finding that, 48 months after the start of the rollout, the service enrollment of beneficiaries who were mailed Tickets late in each rollout was not significantly lower than for those mailed Tickets early in the rollout. Put differently, those mailed Tickets early had a head start on service enrollment relative to those mailed Tickets later, but the latter group caught up after SSA mailed them their Tickets.

We used the service enrollment estimates to project the impact of mailing the Ticket in the first rollout month versus not mailing it at all on cumulative service enrollment as of 12, 24, 36, and 48 months after the rollout start. We project that the total impact on service enrollment at 48 months is considerably larger than the impact at 12 months: 2.3 percentage points in Phase 2 (versus 1.0) and 1.2 percentage points in Phase 3 (versus 0.8).

The analysis provides no consistent evidence of impacts on other outcomes. Some estimates for Phase 2 are suggestive of an impact. Specifically, marginally significant Phase 2 point estimates for STW and NSTW months imply that a 12-month delay in mailing a Ticket increases both the attainment of STW as of 48 months by 0.6 percentage points and the number of NSTW months by an average of 0.07 months. The estimate for STW is about 7 percent of the corresponding percentage for the whole sample at the end of 48 months, and the estimate for NSTW months is about 5 percent of the mean for NSTW months at the end of 48 months.

Although the Phase 2 estimates on their own suggest that TTW had an impact on STW and NSTW months, there are substantial reasons to believe that the results are simply due to chance. The comparison of the Phase 3 impact estimates for STW and NSTW months with those for Phase 2 is particularly important because of the larger Phase 3 sample and a rollout period that was equal in length to the Phase 2 rollout period. The Phase 3 estimate for the impact on STW at 48 months is not significant; it is also only one-third the size of the Phase 2 estimate. The Phase 3 estimate of the impact on NSTW months is comparable in magnitude to the Phase 2 estimate, *but in the opposite direction*. Therefore, we can as easily interpret the Phase 3 estimates as evidence that the Ticket mailing *reduced* NSTW months as we can interpret the Phase 2 estimates as evidence that the Ticket mailing *increased* NSTW months. It is difficult to understand why comparable impacts on service enrollment in the two samples would translate into such different impacts on NSTW months.

Results for TWP completion also undermine the conclusion that the Phase 2 estimates for STW and NSTW months reflect real impacts. We found almost no relationship between duration to MM and TWP completion in Phase 2 and marginally significant evidence of a negative impact of duration to MM on TWP completion in Phase 3. This is the opposite of what we would expect if the Phase 2 estimates for STW and NSTW months reflected real impacts in Phase 2, and if the Phase 3 estimates for the same variables indicated an absence of real impacts in Phase 3.

The projections for total impacts on outcome variables other than service enrollment also provide no evidence of positive total impacts. Most notably, projections of cumulative impacts of TTW on STW and number of NSTW months for Phases 2 and 3 were not significantly different from zero as of any of the four observation points.

In summary, the consistency of the findings on service enrollment at 12 months convinces us that those estimates reflect real impacts. Symmetrically, the inconsistency of the findings on TWP start, TWP completion, STW and NSTW months across these same samples implies that there is no evidence of positive impacts for these outcomes.

The Hypothesis That TTW Was Self-Financing by 2007 Cannot Be Rejected

Based on previous TTW evaluation reports and reasonable additional assumptions, it is plausible that TTW, as initially implemented, would be self financing (that is, that it would pay for itself through benefit reductions) even if its impacts on STW and the subsequent number of NSTW months were very small—possibly as small as 5 percent based on 2007 data. We used the projections of impacts on STW and NSTW months at 48 months to test the hypothesis that TTW was self-financing in 2007 versus the alternative that it was not.

We are not able to confidently reject the self-financing hypothesis on the basis of the projected impacts for either STW or NSTW. Thus, although the results overall are consistent with no impact, and we are quite confident that the estimates are not biased, the statistical power of the methodology is insufficient to definitively discriminate between "no effect" and "smallest effect consistent with self-financing." Consequently, we are unable to confidently conclude that the program was not self-financing prior to the 2008 regulatory change. Nonetheless, our assessment is that the results are more consistent with the hypothesis of no impact than they are with the hypothesis that TTW was self-financing in 2007.

Conclusion

The methodological approach used in this report to estimate the initial impacts of TTW on key outcomes is much stronger than the methodology that we were able to apply in earlier reports. We have considerable confidence that the methodology provides unbiased estimates and that it can detect impacts that are quite small. The evidence of positive impacts on service enrollment is consistent with the evidence from earlier analyses, and quite strong.

We conclude, however, that there is no evidence of positive impacts on two key variables, STW and NSTW months. Although we cannot confidently rule out the possibility that there were positive impacts, we must conclude that the weight of the evidence is more consistent with "no impacts" on these outcomes than the conclusion that TTW was self-financing under the original regulations.

It is important to keep in mind that these estimates are for TTW under the original regulations. The 2008 changes to the regulations clearly stimulated provider interest and beneficiary participation

in TTW. Those changes might have had a positive impact on STW and NSTW months. Unfortunately, it appears impossible to rigorously measure any such impact because the regulations were implemented nationally, without a test, and also because implementation of the new regulations coincided with the trough of a major recession in the U.S.

The new analysis provides a lesson for SSA and other agencies when, in the future, they are asked to make a significant change to a large national or state program—including significant future changes to TTW. Inasmuch as such a change often requires a lengthy rollout period, agencies should consider the knowledge that might be gained by implementing a rollout in which participants are randomly assigned to an implementation month over a period of 12 months or so. This approach has its limits, however; it will not necessarily have sufficient statistical power to identify policy-relevant impacts if such impacts are very small. Power can be increased through enhancements that make this approach more like the approach that would be best from a purely methodological perspective: a randomized control trial in which those likely to be affected by the change are randomly assigned to either a treatment group or a control group, and mean outcomes for the two groups are compared over a period of sufficient length to estimate impacts of interest to the agency.

This page has been left blank for double-sided copying.

I. INTRODUCTION

Social Security Disability (SSD) benefits are available to workers who experience long-lasting medical impairments that prevent work at a substantial level (disabled workers), as well as to Disabled Adult Children (DAC) and Disabled Widow(er)s of other Social Security beneficiaries.[1] In 2011, more than 9.8 million people received SSD benefits.[2]

Many SSD beneficiaries are able and willing to work at some level; most of those who work earn little enough that they do not lose their benefits. Recognizing this, the Ticket to Work and Work Improvement Incentives Act of 1999 (Ticket Act) put into place a number of new policies and programs designed to encourage beneficiaries' return-to-work efforts. The policies and programs include initiatives that provide beneficiaries with information about how work affects their benefits, offer them more options for accessing employment services, allow them to return more easily to the disability rolls after unsuccessful work attempts, and facilitate the processing of earnings information by SSA staff.

The leading initiative authorized by the Ticket Act is the Ticket to Work (TTW) program. Under TTW, the Social Security Administration (SSA) mails each eligible disability program beneficiary a "Ticket" that he or she could assign to either a state vocational rehabilitation agency (SVRA) or to a prequalified local rehabilitation service provider, called an employment network (EN), in exchange for employment placement, job training, and other services. SSA promised to pay the provider on the basis of earnings and benefit outcomes for the beneficiary. TTW was designed to expand the service options available to beneficiaries and create greater incentives for providers to help beneficiaries earn enough to forgo benefits. Prior to TTW, SSA essentially paid SVRAs only for providing services, and only under a cost-reimbursement payment system that had earnings incentives, but did not offer incentives for reducing benefits to SSD beneficiaries.

TTW was rolled out in three phases. A first set of states completed the TTW rollout in 2002 (Phase 1), a second set in 2003 (Phase 2), and a final set in 2004 (Phase 3). In July 2008, SSA significantly changed the regulations governing TTW to attract more providers and reflect a more flexible return-to-work concept; hereafter, we call the pre-2008 program the "original" program

In this report, we present results from a new analysis of the impact of the introduction of the original TTW program using longitudinal administrative data for young, new SSD beneficiaries. Although it would be more interesting to evaluate the impact of TTW under the new regulations, that is not feasible, because the regulations were changed nationally in the midst of a major recession. Impacts under the original regulations are nonetheless interesting. Previous attempts to

[1] DAC receive benefits on the basis of a parent's entitlement as a "primary beneficiary"—a parent who is a disabled worker, retirement beneficiary, or deceased worker. The DAC must be deemed unable to work as of the age of 22 under the same medical criteria applied to disabled workers, he or she is not entitled to benefits until the parent is entitled. Each disabled widow(er) beneficiary (DWB) receives benefits on the basis of the entitlement of a deceased spouse; the DWB must be at least 50 years old as well as meet the same medical criteria as disabled workers. DAC and DWB benefits are paid out of the Social Security Disability Insurance (SSDI) Trust Fund if the primary beneficiary is a disabled worker, or out of the Old Age and Survivors Insurance (OASI) Trust Fund if the primary beneficiary is a retiree or deceased. See SSA (2012) for further details.

[2] Because the analysis presented in this report includes SSDI disabled worker beneficiaries as well as DAC and DWB, and benefits for most of the latter are not paid from the SSDI Trust Fund, we use SSD to encompass all three groups.

estimate impacts on earnings and benefits were, in essence, based on annual trends in differences for mean outcomes across the three phases (Thornton et al. 2007; Stapleton et al. 2008). Results were inconclusive, because methodological issues made it impossible to discriminate between potentially very small, yet important impacts of TTW and pre-existing trends in the differences across phases for earnings and benefit outcomes.

The impact estimates in this report are based on an approach developed under the guidance of the project's Technical Evaluation Support Group (TESG) and is methodologically stronger in multiple respects. First, the passage of time has allowed us to observe the behavior of those who were mailed a ticket early in the rollout period for a much longer period; that's critical because it may take a long time for a beneficiary to return to work and earn enough to forego benefits, and even longer before such changes are fully reflected in administrative data. Second, we take advantage of two recently developed monthly outcome measures—a monthly indicator for nonpayment status following benefit suspense or termination because of work. The previous analysis had to rely on an annual measure of earnings, whereas tickets were distributed on a monthly basis. Finally, rather than rely on the across-phase comparisons of outcomes to identify impacts, we utilize the essentially random, exogenous variation in Ticket mail month during the rollout within each phase to measure the impacts of TTW on monthly outcome variables.

We also decided to focus on the subgroup of beneficiaries for which we expect impacts to be easiest to detect: young (ages 18 to 39 at award), new SSDI-only beneficiaries—that is, those not also receiving Supplemental Security Income (SSI). The earlier impact analysis concluded that impacts on service enrollment per beneficiary were largest for SSD-only beneficiaries ages 18 to 39 at the start of the TTW rollout.[3] More recent research has also shown that young, new SSD beneficiaries return to work and eventually forego benefits for work at a much higher rate than others (Liu and Stapleton 2011). We assumed that if we found no substantial evidence of impacts on key outcomes for this group, we could be reasonably confident that there would be no evidence of impacts if we applied the same approach to other groups. If instead we found evidence of positive impacts for this group, we could extend the methodology to see if the same approach produced positive evidence for other groups. The focus on new beneficiaries also allowed us to exclude those who received awards before July 1999, when there was a significant change in how earnings affect benefits.

The five outcome variables are enrollment for employment services with an SVRA or EN, starting the trial work period (TWP), completing the TWP, first month of benefit suspense or termination for work (STW),[4] and the number of months in nonpayment status following STW

[3] In the previous analyses, we estimated that TTW increased service enrollment for young SSD-only beneficiaries as of the end of the year after the rollout year by 1.4 percentage points (Stapleton et al. 2008, Exhibit XII.5). The estimated percentage-point impacts on service enrolment for young concurrent and SSI-only beneficiaries are slightly lower (1.3 and 1.1, respectively). Estimates for older beneficiaries in all age-program groups are considerably lower; the highest estimate for the older groups is 0.8 percentage points, for SSD-only beneficiaries aged 40 to 49. All of these estimates are significant at the 0.01 level.

[4] During the TWP, SSD beneficiaries are permitted to work and earn at any level without loss of benefits, provided that they continue to meet the medical eligibility requirements. The TWP consists of 9 months, which need not be consecutive—any 9 months in a 60-month rolling window are counted. After completing the TWP, beneficiaries enter an extended period of eligibility (EPE). Except for a 3-month grace period, individuals who earn more than engage in substantial gainful activity (SGA) in any of the next 36 months have their benefits suspended for that month. The beneficiary is entitled to full benefits during any month of this period when he or she is not engaged in SGA, provided that benefits have not been terminated for medical recovery or some other reason. After 36 months, SSD benefits are terminated in the first month of SGA, or in the first months after use of any remaining grace period months.

before returning to current-pay status, attainment of the full retirement age (FRA), or death (NSTW months).[5, 6] As will be seen, circumstances are very favorable to this approach in Phases 2 and 3: the sample sizes in the Phase 2 and 3 states are very large, and the variation in duration from rollout start to Ticket mail month is substantial. Circumstances are much less favorable for the analysis in Phase 1, where the planned rollout was compressed and where we had to split the sample into two pieces because an operational issue led to different rollout schedules for New York (NY) and the rest of the Phase 1 states.

In the remainder of this chapter, we provide context for the report, describe the research questions we have addressed, summarize the key findings, and provide an overview of the rest of the report.

A. Background

Findings from previous analysis of early impacts of the TTW program suggest that TTW increased enrollment in employment services, and that there were larger impacts for younger beneficiaries than for older beneficiaries, and little variation in impacts by title–SSD-only, SSI-only, and concurrent (Thornton et al. 2007; Stapleton et al. 2008). These impacts were estimated by comparing service enrollment across the three phases of the rollout; in essence, the states in the later phases were used as comparison groups for the states in the earlier phases.

The same approach produced inconclusive evidence on whether TTW affected beneficiary earnings and benefits during its first two years, however. If TTW had any success in increasing beneficiary earnings or reducing receipt of benefits, those effects were masked by two other factors: (1) differences across rollout phases in employment and benefit-receipt trends that pre-dated the TTW program, and (2) underlying variation in beneficiary outcomes across states and over time (Thornton et al. 2007). The necessity of using an intent-to-treat approach for a program with very low participation rates (1 to 2 percent), nonrandom variation in the effects of external factors across phases, and annual outcome measures for earnings and benefits made this approach all the more challenging. The findings left open the possibility that there were earnings and benefit impacts of sufficient size to be of interest for policy purposes, but if there were such impacts, they were obscured by the limitations of the methodology.

Given the ambiguity about impacts of TTW on earnings and benefits, we explored a number of alternative strategies for estimating impacts of TTW with the TESG. The TESG recommended implementing some of the alternatives. The recommended strategies include (1) use of the National Beneficiary Survey (NBS) as if it were a random-outreach intervention, (2) estimation of impacts of Ticket mailing on duration from award to key outcomes using exogenous variation in timing of mailing relative to award month from all available sources (namely, variation in the timing of SSD award, across-phase variation in the rollout start, and random within-phase variation in Ticket mail month), and (3) estimation of impacts of duration from the start of each phase's rollout to the month of Ticket mailing using random within-phase variation in mail month. In addition, they recommended focusing on recent young SSD-only awardees and estimating impacts on four monthly outcomes—enrollment in employment services, beginning of the TWP, completion of the

[5] "Current-pay" status means that the individual is eligible for a cash payment for the current month.

[6] In other recent reports submitted to SSA under the TTW evaluation, STW is referred to as nonpayment status following suspension or termination for work (NSTW). In this report we use STW to avoid confusion with NSTW months.

TWP, and STW. They also recommended extending the analysis of impacts beyond two years after Ticket mailing, as it might take longer than two years for impacts of TTW to materialize on outcomes other than service enrollment.

Use of the NBS as a random-outreach intervention. The NBS is a cross-sectional survey of a nationally representative, randomly selected sample of working-age Social Security disability beneficiaries. The interviewer queried each respondent about his or her awareness and knowledge of the TTW program, and, in doing so, delivered information on the existence and goals of TTW. As a result, NBS interview attempts (successful or not) can be considered a randomized-outreach intervention—one that might affect the likelihood of Ticket assignment and, consequently, other outcomes. The identification strategy for estimating impact of Ticket use on employment-related outcomes and benefit receipt involved the NBS interview attempt as the instrumental variable for Ticket assignment. Whalen et al. (2011) determined, however, that the NBS interview attempt is a poor instrument for estimating the impact of Ticket use on employment outcomes and benefit receipt, primarily because it is not a statistically significant predictor of Ticket use.

Use of exogenous variation in timing of mailing relative to award month from all available sources. This strategy for estimating impacts of TTW involves applying a difference-in-differences approach that takes advantage of all sources of variation in the duration from award month to the Ticket mail month: variation in the month of SSD award, across-phase variation in the TTW rollout schedule, and random within-phase variation in the Ticket mail month. Hereafter, we refer to this analysis as "all-variation (AV) analysis." To succeed, such analysis must control for factors that might be correlated with these sources of variation and have their own impacts on outcomes.

The initial AV findings for duration from award to each outcome appeared promising, but the findings were not robust to a key specification test. Using data for all phases, we found positive and significant impacts of Ticket mailing on service enrollment, TWP start, TWP completion, and STW in the first or second month after award. However, when we estimated within-phase models, using only variation due to award month and within-phase variation in mail month, we found evidence of impacts on duration to service enrollment in only two of the three phases, and did not find substantial evidence of impacts on duration to TWP start, TWP completion, or STW. We think that the results using data for all phases reflect specification bias—that is, the estimated model was unable to fully capture external factors, possibly including the business cycle. In addition, estimates from the within-phase models are less precise than those from the model using data for all phases because of smaller sample sizes. Further details on the all-variation analysis and its shortcomings are available in a memorandum (Stapleton and Mamun 2012).

Use of random within-state variation in mail month alone to identify impacts. This strategy for estimating impacts of TTW involves applying the within-state random variation in Ticket mail month resulting from the use of the terminal digit of the beneficiary's Social Security number (SSN) to determine each beneficiary's intended mail month.[7] The analysis using the random variation in intended mail month estimates the impacts of duration from the start of the rollout to the mail month on beneficiary outcomes over a 48-month period starting with the first rollout month. This report focuses on analysis using this strategy.

[7] Because the last four digits (called the "serial numbers") of SSNs were historically assigned in sequence (as described in Barron and Bamberger 1982), the SSNs are correlated with age. Therefore, the Ticket mailing date is considered random conditional on age.

We focus on recent young SSD awardees—beneficiaries who were ages 18 to 39 at the time they entered the SSD program. The focus on this group is attractive because: (A) earlier analyses indicate that introduction of TTW had a larger impact on their service enrollment than on that of other SSD beneficiaries and SSI recipients (Thornton et al. 2007); and (B) new evidence that young new SSD awardees are much more likely than others to enroll for services, work, and eventually earn enough to leave cash benefits for work (Liu and Stapleton 2011; Mamun et al. 2011; Stapleton et al. 2008). Our use of SSD-only awardees, rather than SSI awardees, reflects the availability of data on cohorts of new SSD-only awardees at the time we began this analysis.[8] We elected to not include concurrent awardees because many young concurrent beneficiaries receive their SSI award first, which created technical problems that could be solved only with very considerable effort. In addition, because of the differences in programmatic rules for the SSD and SSI programs, the concurrent beneficiaries faced different work incentives than SSD-only beneficiaries, so pooling concurrent and SSD-only beneficiaries would complicate the analysis.[9] By focusing on young, new SSD beneficiaries we are able to establish whether the original TTW program had any impacts on a subgroup that has the most promise. Given the ambiguity of the earlier findings, the use of this subgroup of beneficiaries allows us to assess the prospect of finding favorable impacts for the full SSA disability beneficiary population. Absence of policy relevant impacts on STW and NSTW months for this subgroup would give us greater confidence that there were no policy relevant impacts for the larger population either.

For this report, we followed individuals for a longer period after the start of the rollout in their state than was possible for the early TTW evaluation reports—in some cases up to 48 months. To improve the precision of the estimates, we used linear models that control for beneficiary characteristics, including pre-TTW benefits and pre-SSD earnings. We also used instrumental variables method to eliminate the possible bias resulting from the fact that, for a small minority of cases, SSA did not mail the Ticket in the intended mail month. Technically, we used the intended mail months as instruments for the actual mail months.

B. Research Questions Addressed

We are able to directly address the following primary research questions with the available data. Each question concerns the impact of duration from the month before the rollout start in the beneficiary's state to the month in which SSA mailed a Ticket to the beneficiary (the beneficiary's "mail month") on outcomes over the 48 months after the rollout start.

- Was each of four event outcomes less likely to be attained as of 12, 24, 36, and 48 months after rollout start the longer the duration from rollout start to mail month?

[8] Data on new concurrent beneficiaries are also available, but many of these beneficiaries received SSI before they became eligible for SSD. To expand the analysis to all new SSD or SSI awardees, we would have to develop data for cohorts of new "disability awardees" based on the first month of SSD or SSI benefit payments. Note that some of SSD beneficiaries in the analysis sample may have become SSI recipients during the observation period.

[9] For instance, concurrent beneficiaries do not necessarily lose all of their benefits when they engage in substantial gainful activity (SGA) after completion of the TWP and grace period months, as some may continue to receive SSI under Section 1619(a)—especially those with a very low SSD benefit. Others, however, may lose their SSI cash payment even before they engage SGA—especially those with a relatively high SSD benefit. Hence, for concurrent beneficiaries, we would have to consider STW from both SSD and SSI, and the timing of STW for SSD and SSI might be different.

- Was the number of months in nonpayment status following suspension or termination for work (NSTW months) as of 12, 24, 36, and 48 months after rollout start smaller the longer the duration from rollout start to mail month?

We initially addressed these questions for each of the three phases separately. Further, because the NY rollout followed a different schedule than the rollout for the rest of the Phase 1 states, we split the first phase sample into an NY sample and a Phase 1 Except NY sample. This structure recognizes that the analysis focuses on the 48 months after the rollout starts and uses random variation of mail month within the phase to identify the impacts of duration to mail month. Because the rollouts for Phases 2 and 3 followed exactly the same pattern separated by 12 months, we also estimated models based on pooled data from those two phases, and tested whether the results from the two phases differ significantly from each other.

The findings from the analysis for the primary research questions can potentially be used to indirectly address two research questions of more immediate interest to policymakers:

- What was the impact on each outcome of mailing the Ticket immediately versus not mailing it at all as of 12, 24, 36, and 48 months later; that is, what was the total impact on each outcome as of each of these points in time?
- For each beneficiary induced to enroll for services by Ticket mailings, what was the mean increase in NSTW months as of 12, 24, 36, and 48 months later?

The first question cannot be answered directly, because SSA eventually mailed a Ticket to essentially everyone in the sample. If we find, however, that the duration from rollout start to Ticket mailing had impacts on beneficiary outcomes, it would necessarily be the case that mailing Tickets had impacts relative to never mailing Tickets. In fact, under a reasonable set of assumptions it is possible to convert the estimates of the impacts of duration to mail month to estimates of the total impacts of mailing the Ticket.

The second question concerns the issue of the extent to which providing employment services to beneficiaries under the Ticket financing mechanism leads to the desired outcome. The government is using TTW to invest in services in a manner that is anticipated to have a return: less reliance on income support from SSD and SSI. If we assume that any impacts of Ticket mailing on NSTW months are the result of induced enrollment for services, we convert the estimates of total impacts on service enrollment and NSTW months to impacts on NSTW months per unit impact on service enrollment. Of course it is possible that total impacts on NSTW month are partly due to other effects that mailing a Ticket have on beneficiary behavior; for instance a beneficiary might interpret the receipt of a Ticket as a message from the government that they should be trying to work and leave the rolls.

C. Key Findings

We find clear evidence that the mailing of Tickets during the rollout period did increase service enrollment. The estimated magnitude of the impact on service enrollment for the recent young SSD-only awardees is very similar to previous estimates of impacts for all young SSD-only beneficiaries. However, we found no consistent evidence that this impact translated to an increase in the number of months in which beneficiaries did not receive benefits following suspension or termination for work. Key estimates for this outcome from the Phase 2 rollout are marginally significant and consistent with a modest impact, but Phase 2 findings for TWP outcomes and findings from

Phase 3 do not reinforce the Phase 2 findings. It seems likely that the marginally significant Phase 2 estimates are simply due to chance.

Even though we do not find positive evidence of impacts on STW and NSTW months, the estimates do not allow us to rule out the possibility that TTW was self-financing in 2007. Based on analyses presented in earlier reports, self-financing would require an increase in these outcomes of as little as 5 percent. Although the results overall are consistent with no impact, and we are quite confident that the estimates are not biased, the statistical power of the methodology is insufficient to definitively discriminate between "no effect" and "smallest effect consistent with self-financing." Nonetheless, the weight of the evidence is more consistent with "no impacts" on these outcomes than the conclusion that TTW was self-financing in 2007.

It is important to keep in mind that the impact estimates presented in this report are for TTW under the original regulations. The 2008 changes to the regulations clearly stimulated provider interest, beneficiary participation in TTW, and the number of NSTW months accruing to participants (Schimmel et. 2013). As they point out, however, it is not feasible to distinguish between the impacts of participation on NSTW months and the effects of an increase in TTW participation by those who would have accrued NSTW months even if they had not used their tickets. That is partly because the regulations were implemented nationally, without a test, and partly because any impacts of the new regulations are confounded with the impacts of a major recession that reached its trough just as the new regulations were being implemented.

D. Organization of the Report

In Chapter II, we summarize the selection of the analytic samples and variables, and describe how we estimated the impacts of duration to mail month on the outcome variables. In Chapter III, we present the main findings from the analysis along with projections of total impacts of Ticket mailing on the outcome variables. We summarize the findings and draw conclusions in Chapter IV. We present the details of the data, methods, and findings in Appendices A through D.

This page has been left blank for double-sided copying.

II. DATA AND METHODS

In this chapter, we briefly describe the data and discuss the methods we applied to conduct the impact analysis. We present the criteria for inclusion in our analytical samples, and describe the sample. We also assess variation in the business cycle over the sample period and present the unemployment rate measures used in later chapters to assess the effect of the economy on Ticket impacts. We describe the strategy used to estimate impacts and provide a synopsis of other methods that we applied to a preliminary data set and an explanation of why we elected to proceed with the analysis presented here. More detailed description of the data, including sample selection, evidence of statistical equivalence of the intended mail month samples within each phase, and descriptive statistics on the outcome measures are presented in Appendix A. Further exposition of the methods appears in Appendix B.

A. Ticket Research File

We used data from the 2007 TRF (TRF07). The TRF is a set of analytic administrative data files constructed for the TTW evaluation. The TRF07 files contain current and historical information on more than 22 million SSD beneficiaries or SSI recipients who received a benefit in at least one month from January 1996 through December 2007 (Hildebrand et al. 2009).[10] For the purpose of this study, we constructed annual cohort files for those awarded benefits from 1999 through 2003.[11] Cohort assignment is based on the month that SSA first paid a benefit to the awardee. Although it is possible for an individual to have multiple entitlements, he or she is assigned to just one cohort based on the year that corresponds to the individual's *first* payment.[12] All analyses were conducted using pooled data from multiple cohorts.

B. Analytic Samples

1. Sample Selection

The sample of interest includes beneficiaries who entered the SSD rolls from July 1999 through October 2003. For the analysis, we followed each beneficiary for 48 months starting with the first month of the rollout in the beneficiary's state. As the Phase 3 rollout started in November 2003, the last month in the sample is October 2007. We limit the analysis to this period because of factors external to the introduction of TTW. We started with July 1999 SSD awardees because this is the month in which the non-blind substantial gainful activity (SGA) level was increased from $500 to $700. We end the follow-up period in 2007 because of the severe recession that started in the last quarter of 2007 and because SSA made substantial changes to TTW regulations in 2008 that may have affected beneficiary outcomes in 2008 and later.

[10] Extracts from several Social Security administrative files were merged to create the TRF, including the Master Beneficiary Record, Supplemental Security Record, Numerical Identification System (Numident) file, the 831 and 832/33 Disability files, the Disability Control File, monthly snapshot files, and files from the payment history update system.

[11] These annual cohort files are an extension of those created by Liu and Stapleton (2011).

[12] The first payment month (that is, the award month) is that in which the first payment was actually made, which is usually after the first month for which the beneficiary is entitled to a benefit (that is, the entitlement month). The latter is often used in SSA's statistics to classify beneficiaries by entry year (for example, SSA 2009). We use the award month instead because our focus is on the activities of beneficiaries once they become informed of their award and are entitled to use the DI work incentives.

The analysis samples consist of young (ages 18 to 39 at award) SSD-only awardees who were first paid SSD benefits no earlier than July 1999 and were selected for the initial rollout of the TTW program on one of three Ticket selection dates: January 12, 2002 (Phase 1), October 26, 2002 (Phase 2), or October 18, 2003 (Phase 3).[13] It was SSA's intent to mail Tickets to every beneficiary in these samples during a subsequent rollout month (hereafter, the "intended mail month" [IMM]), to be determined by the terminal digit of the beneficiary's SSN. Exhibit II.1 provides sample size for each of the four samples used in the analysis and shows the number of beneficiaries assigned to each IMM. As will be seen, SSA mailed the vast majority of these Tickets on the IMM. For each phase, we treat the samples defined by IMM (hereafter, the "IMM samples") as randomly assigned samples of those included on the phase's selection date. Further details on sample selection, sample sizes by IMM, and statistical equivalence of IMM samples within each phase are presented in Appendix A.

Although SSA actually mailed Tickets on the IMM for most beneficiaries, for a small fraction the actual mail month (MM) did not correspond to the IMM. The TRF records include the actual mail date, making it possible to determine the MM. Across the four phase-samples, in 93 to 99 percent of the cases the MM is the IMM. Although the fraction of Tickets mailed on the IMM was very high in each month of the rollout, it did decline in successive months. One reason for the decline is a provision of the regulations called "Ticket on demand," as beneficiaries assigned to a later IMM had more time to request a Ticket on their own. In addition, as the rollout progressed, SSA identified some beneficiaries who had died or were no longer in current-pay status, and consequently did not mail these beneficiaries their Tickets. Mailing a Ticket on demand, mortality, and loss of current pay status for some other reason is likely predictive of beneficiary outcomes of interest. This poses challenges for estimating impacts of duration to Ticket mailing on beneficiary outcomes. We were able to rigorously address this challenge, as described in Section C of this chapter.

As shown in Appendix A, the IMM samples are not statistically equivalent in terms of the means for beneficiary characteristics observed in or before the selection month, but differences are substantively very small. The significant differences likely reflect that the method SSA used to assign IMM was not purely random, particularly with respect to age and factors associated with age, coupled with sample sizes that make even non-substantive differences statistically significant. We used standard statistical methods to adjust for the small differences in observed characteristics.

2. The Economy

The Ticket rollout occurred during the economic expansion following the 2001 recession, but the strength and timing of the recovery varied across states and rollout phases. By examining monthly national unemployment rates from the first rollout month in each phase through the 48th month (the analysis period for each rollout), we find that all three analysis periods are predominantly periods of recovery, but the timing of recovery varies substantially across phases.

Of course, all beneficiaries within each phase were subject to the same national economic circumstances during the analysis period for their phase, so there is no reason to be concerned about bias in the impact estimates from within-phase correlation between a beneficiary's mail month and

[13] SSA determined all beneficiaries who were eligible to receive a Ticket and who resided within the phase's states as of the phase's selection month. Almost all SSD beneficiaries and SSI recipients over age 18 were eligible; the main exceptions were (1) new beneficiaries with a status of medical improvement expected (MIE) who had not yet had their first medical continuing disability review (medical CDR) and (2) child SSI recipients who had reached age 18 and were waiting for redetermination as adults.

Exhibit II.1. IMM Sample Sizes by Phase

	Phase 1 NY			Phase 1 Except NY			Phase 2			Phase 3		
Rollout Month	Intended Mail Month	Calendar Month	N	Intended Mail Month	Calendar Month	N	Intended Mail Month	Calendar Month	N	Intended Mail Month	Calendar Month	N
1	1	Feb-02	1,148	1	Feb-02	4,282	10	Nov-02	7,573	21	Nov-03	11,531
2												
3				3	Apr-02	8,644	12	Jan-03	7,733	23	Jan-04	11,328
4	4	May-02	1,163	4	May-02	12,960	13	Feb-03	7,679	24	Feb-04	11,539
5				5	Jun-02	17,194	14	Mar-03	7,745	25	Mar-04	11,569
6	6	Jul-02	2,438				15	Apr-03	7,743	26	Apr-04	11,523
7	7	Aug-02	2,408				16	May-03	7,778	27	May-04	11,533
8	8	Sep-02	2,452				17	Jun-03	7,598	28	Jun-04	11,434
9	9	Oct-02	2,414				18	Jul-03	7,826	29	Jul-04	11,519
10							19	Aug-03	7,681	30	Aug-04	11,253
11							20	Sep-03	7,805	31	Sep-04	11,428
Total			**12,023**			**43,080**			**77,161**			**114,657**

Notes: Because the TTW rollout in NY followed a different schedule than the rollout for the rest of the Phase 1 states, we split the first phase sample into an NY sample and a Phase 1 Except NY sample.

the economic conditions experienced by the beneficiary. It is at least arguable, however, that the impact of Ticket mailing could interact with the state of the economy at the time the Ticket is mailed. For instance, beneficiaries might be more successful finding work on their own, without a Ticket, during a strong expansion than they would be during the trough of the business cycle.

To account for variation in economic conditions at the state level around the beneficiary's IMM, we included two aggregate measures of state-level monthly unemployment rates: first, the average unemployment rate during the six-month period around each beneficiary's IMM (from two months before through three months after the IMM), and second, the change in the state's monthly unemployment rate during the same period.

3. Outcome Measures

The outcome measures are based on the 48 months starting with the first rollout month for the phase (month zero is the pre-rollout month). This period ends in January 2006 for Phase 1, September 2006 for Phase 2, and September 2007 for Phase 3. For each individual in the sample we constructed:

- Four binary "event" variables. We determined when in the 48 months following start of rollout each of the following events occurred, if at all: (1) enrolled for employment services (assigned their Ticket to an EN or were determined eligible for services by an SVRA); (2) completed their first TWP month; (3) completed their last TWP month; and (4) had their benefits suspended or terminated for work. In the analysis of whether an event has occurred as of a specified rollout month (month 12, 24, 36, or 48), we define a binary for each event that is equal to one if the event occurred after the rollout start and before that month, and zero otherwise.
- NSTW months, a count of the number of months in nonpayment status following STW that occurred during the 48-month period. NSTW months include all months after benefits are suspended or terminated for work until the first of the following events occurs: (1) return to current-pay status, (2) suspension or termination for some other reason, or (3) the end of the 48-month period. Beneficiaries are not necessarily engaged in SGA during all NSTW months; we know only that they are not receiving benefits.

Means for the outcome variables in the IMM samples as of month 48 are presented in Appendix A.

C. Estimation Approach

The research questions focus on the impact of the duration from the start of the rollout to the Ticket mail month on beneficiary outcomes measured at four 12-month intervals following the rollout start. Our estimation approach uses the fact that IMM were randomly assigned and that the MM and IMM are identical in a very large majority of cases. Here, we first describe a simpler approach that would be appropriate if IMM were assigned in a purely random fashion and the MM and IMM were identical for all beneficiaries in the sample. We then outline modifications to this approach that we implemented to address the limitations of the simpler approach.

If each MM was identical to its IMM, and the IMM was assigned in a purely random fashion, we could have applied the following straightforward and intuitive method within each phase: for any outcome measured as of a specific month following rollout start in each phase, we could compare

the means across the IMM. If duration to IMM in that phase had a negative impact for that outcome as of that month, then we should observe a gradual decline in the means from the first IMM to the last IMM. To estimate the mean change in the outcome for each month that the mailing was delayed, we could then fit a trend line to the means.

This simple approach underlies the estimates we present in this document, but we modified it to account for small departures from the assumptions that underlie the simple approach and to improve the statistical precision of the estimates (that is, to narrow the confidence intervals around the point estimates and increase the power of statistical tests). First, to address the fact that some Tickets were never mailed because benefits were suspended or terminated prior to the IMM, we coded the MM for those observations as if the Tickets were actually mailed on the IMM. We had previously verified that, with almost no exceptions, termination or suspension of benefits had occurred for reasons other than work—most commonly mortality. We used this modification because mailing the Ticket to these beneficiaries during any month of the rollout would almost certainly have had no impact on their employment outcomes, in which case essentially all outcomes would have been the same as those observed had SSA mailed these Tickets in their IMM.

Second, to address the possible effects of small differences in observed means for beneficiary characteristics across IMM within each phase's sample, and to improve precision, we used a regression framework that incorporates these characteristics along with indicators for the MM. In this framework, the coefficients of the MM are interpreted as the mean outcome for beneficiaries mailed Tickets in that month after adjustment for variation in beneficiary characteristics.

Finally, we addressed the fact that a small share of Tickets were mailed in a month other than the IMM—almost always earlier. If we failed to address this issue, the estimates of the impact of duration to MM would be biased because they would be confounded by factors that led to the difference between MM and IMM—most notably that the beneficiary, perhaps at the behest of an SVRA or other EN, requested immediate delivery under the Ticket on demand option. To address this issue, we applied a technique called instrumental variables (IV) in the regression framework. This technique requires the existence of other variables, or instruments, that are correlated with the variables of interest—the MM in our context—and that have no effect on the outcome variable other than through their effect on MM. The IMM are ideal instruments, because they are highly correlated with the MM (if a specific IMM was assigned, the chance that the ticket was actually mailed in that month is much higher than if a different month was assigned), and because the assignment of an IMM has no plausible effect on the outcome except through its effect on MM.

To generate a trend line across the monthly coefficients for each outcome variable, we estimated a variant of the model that required estimated coefficients for the MM to lay on a straight line. This serves the same purpose as fitting a simple trend line to the monthly coefficients, and also allows us to statistically test two hypotheses: (1) that the effect of each month's delay on the outcome variable is constant, regardless of how long the delay has already been (constant marginal effect of the delay), and (2) that the marginal effect is zero. This restricted version of the estimate is used to generate estimates of how a 12-month delay in Ticket mailing affects the outcome variable.

The IV estimates are summarized in Chapter III and detailed results for these models appear in Appendix E as does regression estimates for which we did not use IV to correct for differences between IMM and MM. The estimated marginal effects for a one-month delay using instrumental variables, reported in Chapter III, are slightly larger in magnitude than the estimates obtained from the regressions without use of IV, but the differences are not substantive. Appendix A presents means for key outcome variables by phase and mail month.

Note that results from our analysis reflect the impact of being mailed a Ticket, not the impact of assigning a Ticket. The IV estimates of impacts of duration to MM as well as the linear regression estimates of duration to IMM identify the effect of the opportunity to use a ticket to receive employment services provided by an SVRA or other EN. Although the mailing a ticket could arguably have impacts on earnings and benefits without having and impact on service enrollment, the expectation has always been that any impact on earnings and benefits would be precipitated by an impact on service enrollment.[14] Because only a little more than 2 percent of the beneficiaries who were mailed a Ticket under the original TTW program regulations ever assigned their Ticket to a SVRA or an EN (Stapleton et al. 2008), our ability to identify impacts of Ticket assignment would be very limited.

D. Methods Applied in Preliminary Analysis

We conducted a number of analyses before we were able to identify the Ticket selection date samples used here. We initially approximated the samples by identifying everybody in current pay in the month prior to the phase's rollout (January 2002 for Phase 1, October 2002 for Phase 3, and October 2003 for Phase 3) who lived in one of the respective phase's states when they received their first SSD payment and were also included in SSA's batch mailing at the beginning of one the months in the phase's rollout. These sample selection criteria meant that some with randomly assigned IMM were not included in the sample—those SSA dropped from the mailing list because of suspension or termination; those mailed a Ticket on demand; and those mailed Tickets in a month outside the rollout period. To partially address the selection problem resulting from Tickets not mailed in the IMM, we excluded beneficiaries who had been mailed a Ticket but were deceased as of the last mail month in the rollout period.

The preliminary analyses yielded consistently significant estimates of the impact of duration to MM on service enrollment as of 12 months; evidence consistent with impacts on STW and NSTW months in Phase 2, but conflicting evidence in Phase 1 and, especially, Phase 3; and little evidence of impacts on TWP completion. The evidence from Phase 2 combined with the conflicting nature of Phase 3 evidence led us to obtain the selection date samples in order to address the analysis limitations caused by the fact that SSA did not mail all Tickets as initially intended, especially during the later mail months for each phase. We thought at the time that the variability of estimates for STW and NSTW months across phases, especially between Phases 2 and 3, might be explained by differences in the completeness of the data and the extent to which Tickets were mailed on the IMM. We also made some changes to the specification of the control variables designed to improve the likelihood of detecting small impacts. These methodological improvements addressed the selection issue and improved estimator precisions, as intended. We also tried to assess how impacts varied with respect to the strength of the state labor market around the Ticket mail month. As will be seen, this fine-tuning of the model specification did not change the overall nature of the findings.

[14] The following are two ways that mailing the Ticket might affect earnings and benefits without affecting service enrollment. First, the new payment system incentives and new competition between SVRA and other EN might result in higher earnings and lower benefits for those who would have used employment services anyway. Second, beneficiaries who received a Ticket might have interpreted them as a signal that they should be trying to work and leave the rolls if they could, and some may have successfully done so without the use of employment services.

III. FINDINGS

We present the primary results of the econometric analysis in this chapter. In the following sections, we examine the IV estimates for the impacts of duration from rollout start to MM on the four event variables (service enrollment, TWP start, TWP completion, and STW) and NSTW months.

A. Clear Evidence of Impacts on Service Enrollment

To illustrate how the estimates are interpreted, we first focus on the service enrollment estimates for Phase 2.[15] We then compare the results for service enrollment as of 12 months across phases.

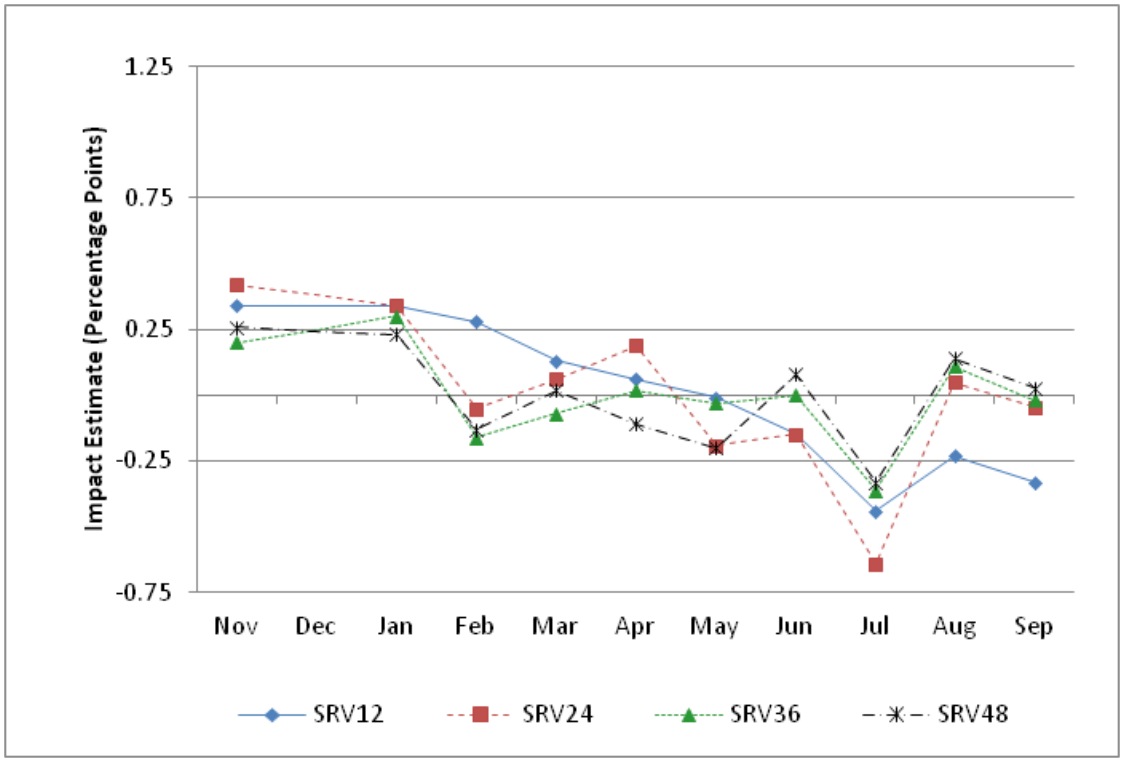
1. Phase 2 Estimates for Impacts on Service Enrollment as of 12, 24, 36, and 48 Months

In Exhibit III.1, we show the point estimates of the MM indicator coefficients for service enrollment as of 12, 24, 36, and 48 months following the start of Ticket rollout in Phase 2. Consider first the estimates at month 12, labeled SRV12. The mean value of the 10 point estimates is zero, by design; each point estimate measures the expected outcome for the sample mailed a Ticket in the corresponding month relative to the overall mean outcome for all those in the Phase 2 sample after adjusting for pre-rollout characteristics and the fact that not all Tickets were mailed on schedule. The point estimate for the first Phase 2 rollout month (November 2002) is 0.34, meaning that, after the adjustments, mean service enrollment as of month 12 following the rollout start for those mailed Tickets in the first rollout month was 0.34 percentage points higher than the overall mean. At the other extreme, the last rollout month (September 2003), the adjusted mean is 0.33 percentage points lower than the overall mean. This implies that, based on these two estimates alone, the effect of a 10-month delay in mailing the Ticket (from month one to month 11 in the rollout), was a reduction in service enrollment as of month 12 of 0.67 (= 0.34 + 0.33) percentage points. Our statistical test of the hypothesis that duration to MM had no effect on this outcome (that is, that the values of the coefficients observed might be due to chance alone) strongly rejected that hypothesis. Any estimate reflects random variation to some degree, but there is little doubt that a negative impact of duration to MM on service enrollment at 12 months is the underlying source of the distinct pattern observed for this series of estimates.

Estimates at each observation point for months 24, 36, and 48 of the rollout showed a negative relationship between duration and service enrollment, but not as strong as at 12 months. In fact, it appears that the relationship becomes weaker as it progresses from month 12 to 24 to 36 to 48. To illustrate, for 48 months the difference between the point estimates for months one and 11 is only 0.23 compared to 0.67 at month 12. Statistical tests, reported in Appendix C, demonstrate more clearly that the relationship between duration to MM and service enrollment does become progressively less significant. We expected this pattern because as time passes, those mailed Tickets late in the rollout start have more time to catch up to those mailed early in the rollout in terms of service enrollment.

[15] We are using the Phase 2 sample to illustrate our findings because the longer rollout (over a 11-month period) first occurred in Phase 2 and the sample is sufficiently large to detect impacts that are quite small.

Exhibit III.1. Estimated Impacts of Duration to MM on the Likelihood of Service Enrollment at 12, 24, 36, and 48 Months Following Rollout Start in Phase 2



Notes: Instrumental variable estimates of impacts on service enrollment outcomes by rollout month relative to overall sample mean (the values are constrained to sum to 0.0). SRVmm is the set of estimates for impacts on service enrollment as of month mm after rollout start. The horizontal axis is labeled with the calendar months for rollout in Phase 2 (November 2002 to September 2003). The vertical axis is measured in percentage points.

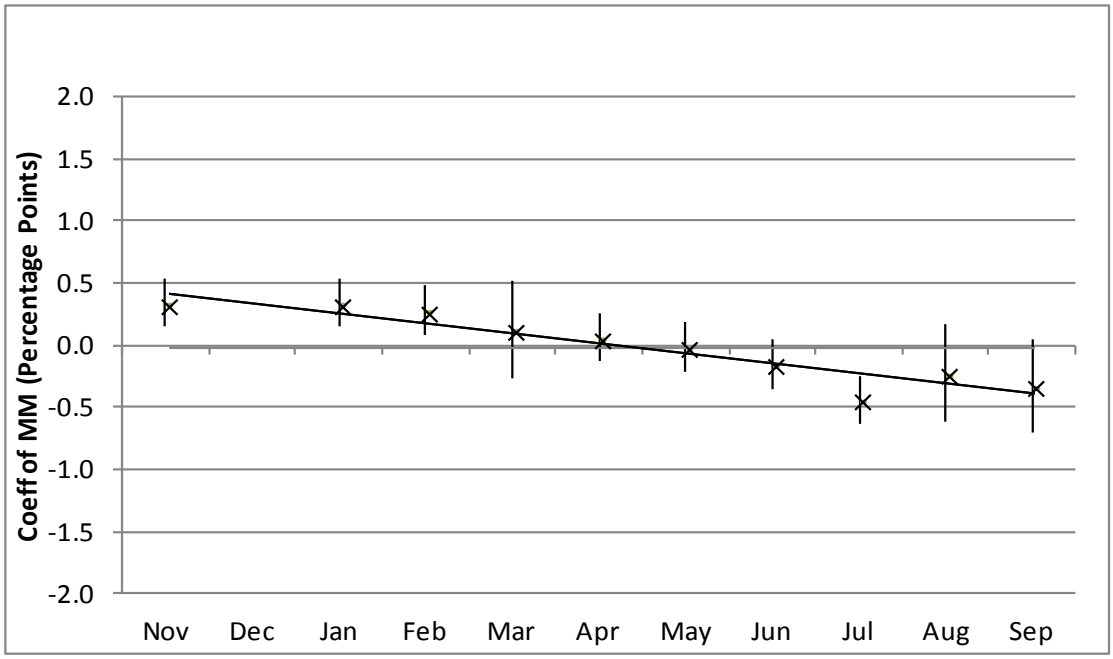
2. Confidence Intervals and Trend Lines for Phase 2 Estimates of Impacts on Service Enrollment at 12 Months

In this section, we focus on the IV estimates of impacts of duration to MM on service enrollment as of 12 months—where we observed the largest impacts, as shown above. The point estimates for this observation month from Exhibit III.1 are repeated in Exhibit III.2 (as Xs). In addition, we have plotted 95 percent confidence intervals around each point estimate (the short vertical line through Xs) as well as the trend line obtained by constraining the IV estimates to fall on a straight line. The confidence intervals are reasonably narrow, and those in the early months lie entirely above zero, while those in the late months lie almost entirely below zero—consistent with the findings from our statistical test. To a first approximation, the point estimates are very close to being in a straight line. In fact, our test of the hypothesis that the true MM coefficients lie on a straight line failed to reject that hypothesis. Thus, it appears that the marginal impact of a one-month delay in mailing Tickets was essentially constant over the 11-month rollout period.

The slope of the line, -0.083 (see Appendix C, Exhibit C.2), implies that each one-month delay in mailing the Ticket reduced the percentage enrolled in services as of month 12 by an estimated 0.083 percentage points. The impact of a 10-month delay is estimated to be 0.83 percentage points—somewhat larger than the 0.67 percentage point estimate we obtained in the discussion of Exhibit III.1 by comparing the estimates of the impacts on service enrollment at month 12 from the

first and last rollout months. Extrapolating to 12 months, the estimated impact of a one-year delay (from month one to month 13) on service enrollment as of month 12 is 1.00 percentage points.

Exhibit III.2. Estimated Impact of Duration to MM on the Percentage Enrolled for Services as of 12 Months Following Rollout Start in Phase 2



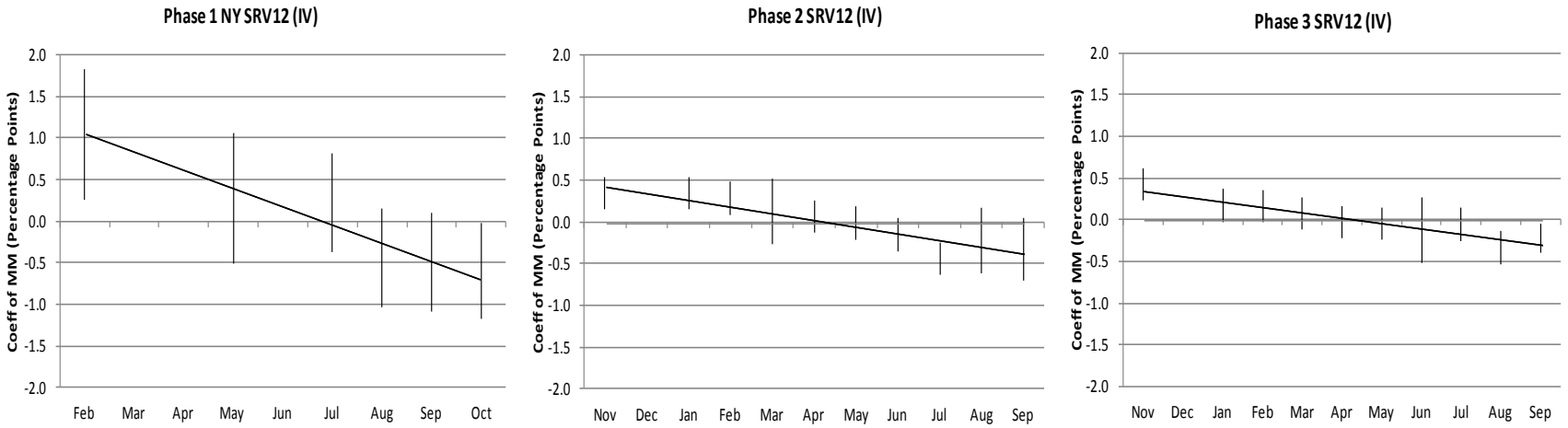
Notes: Instrumental variable estimates and confidence intervals of impacts on service enrollment as of month 12 by rollout month relative to overall sample mean (the values are constrained to sum to 0.0). The trend line was fitted by constraining the instrumental variable coefficients to be on a line. See Appendix C, Exhibit C.2.

Note that a 12-month delay after month one implies that the Ticket was not mailed until after month 12. Hence, this extrapolated estimate can be interpreted as the impact as of month 12 of mailing the Ticket in month one versus not mailing it at all. This interpretation plays an important role in our projections of total impacts on outcomes at 12, 24, 36, and 48 months (versus not mailing the Ticket at all), as reported in Section E below.

3. Comparison of Estimated Impacts of Duration to MM on Service Enrollment at 12 Months Across Phases

In Exhibit III.3, we present graphs with IV estimates of the impacts of duration to MM on service enrollment as of 12 months after rollout start in Phase 1 NY and Phase 3 samples (left and right panel, respectively), along with estimates for the Phase 2 sample (center panel). We omit the Phase 1 Except NY findings here, as well as later—for that sample we find no evidence of a significant relationship between duration to MM and service enrollment at the end of any 12-month interval in the 48 months after the rollout starts. We attribute this finding to very short duration of the rollout for that sample coupled with the relatively small sample size. That is, methodological issues make it unlikely that we would detect substantial impacts with that sample even if they had occurred We also omit the findings for this sample in the presentation of impacts on other outcomes, for the same reason; full results for the sample, which provide no evidence of impacts, are presented in Appendix E, Table E1a.

Exhibit III.3. Estimated Impacts of Duration to MM on the Percentage Enrolled for Services as of 12 Months Following Rollout Start in Phase 1 NY, Phase 2, and Phase 3



Notes: IV estimates of the impact of mailing the Ticket in the rollout month are indicated relative to the average for all rollout months. All estimates are constrained to sum to zero in each phase. The slope of each linear trend line reflects the estimates from the same model with linear restrictions imposed on the coefficients; the coefficients, slope estimates, and test statistics are reported in Appendix C, Exhibit C.2.

As is evident from the exhibits, our qualitative conclusions for the Phase 2 samples apply equally well to the other two samples. In other words, there is statistically significant and consistent evidence of negative impacts of duration to MM on service enrollment at 12 months following rollout start in Phase 1 NY, Phase 2, and Phase 3. Further, it appears that the marginal impact of a one-month delay in each sample is essentially constant throughout the relevant rollout period.

It is also appears from the slopes of the trend lines in the exhibits that the magnitude of the impact of duration to MM diminishes from Phase 1 NY to Phase 2 then again to Phase 3. Extrapolation of the trend lines for the IV estimates in each phase to 12 months provides an estimate of the impact on service enrollment of mailing the Ticket in the first rollout month versus not mailing the Ticket until month 13 or later: 2.5 percentage points for Phase 1 NY, 1.0 percentage point for Phase 2 (reported previously), and 0.8 percentage points for Phase 3.[16] These differences, however, are not statistically significant. Note that the confidence intervals for NY are much wider than those for Phases 2 and 3, reflecting the relatively small sample size and the shorter rollout duration. As a result the substantively large difference between the NY slope and the Phase 2 slope is not statistically significant. The much more modest difference between the Phase 2 and 3 slopes is also not significant, despite much narrower confidence intervals.

We found more limited evidence of impacts at 24, 36, and 48 months following rollout start in the Phase 1 NY and Phase 3 samples. As described above, the instrumental variable estimates for service enrollment as of later observation points (24, 36, and 48 months following rollout start) indicates the those mailed Tickets in later in the Phase 2 rollout substantially caught up to those mailed Tickets early in that rollout by month 48. For Phase 1 NY and for Phase 3, we find no statistically significant effects even earlier—at 24 months, as well as at 36 and 48 months. Taken together, the evidence indicates that those mailed Tickets later generally caught up in terms of service enrollment with those mailed Tickets early during the first and second year following the rollout.

B. Unclear Evidence of Impacts on TWP Start and Completion

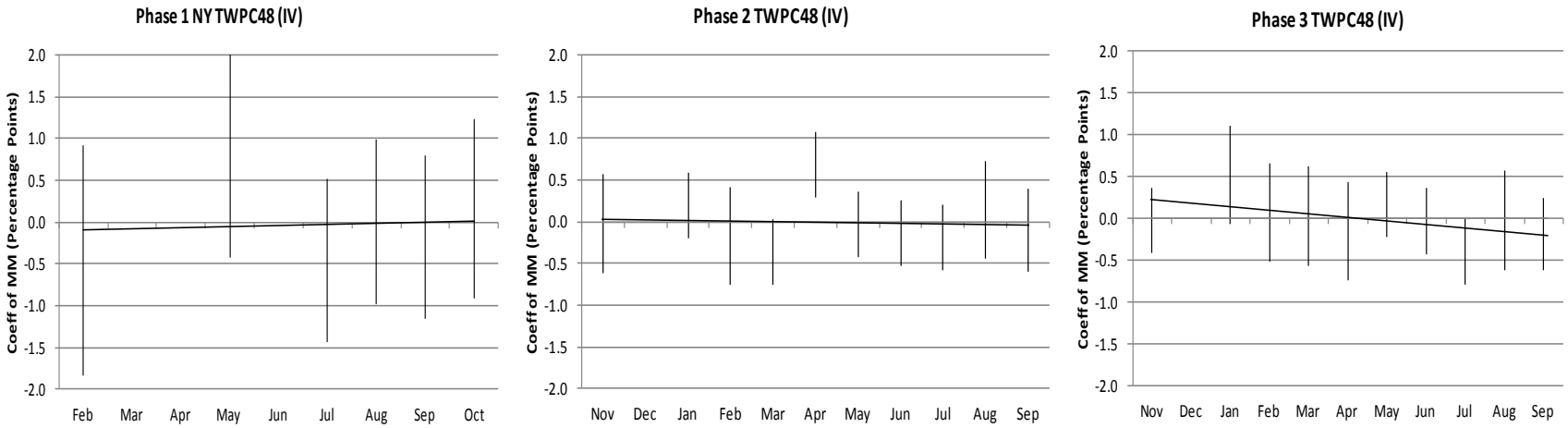
Negative impacts of duration to MM on TWP start or completion would imply that mailing the Ticket induced some beneficiaries to work and earn enough to use TWP months, or to do so sooner than they would otherwise. Almost all estimates of impacts on TWP start—for every sample and every observation point—were statistically insignificant, however, and we do not consider them further here (all estimates are reported in Appendix E, Table E1b). Estimates of impacts on TWP completion are more often significant, but present a mixed picture.[17]

For TWP completion, Exhibit III.4 plots the instrumental variable estimates for impacts on the likelihood of TWP completion at 48 months after the start of rollout, along with their 95 percent confidence intervals and estimated trend lines. Estimates for TWP completion at 12, 24 and 36 months appear in Appendix E, Table E1c, and are no stronger in terms of evidence of impacts than those at month 48.

[16] These values were obtained by multiplying the slopes of the respective fitted lines by 12 (to extrapolate the estimated impacts for 12 months).

[17] One factor that might influence estimates for TWP start is that it is underreported in the administrative data. Typically we have found that about 20 percent of those with a TWP completion date do not have a TWP start date. This likely reflects the fact that the TWP start date has no immediate consequence for benefits due, whereas a TWP completion date is a critical marker for purposes of determining future benefit payments.

Exhibit III.4. Estimated Impacts of Duration to MM on TWP Completion as of 48 Months Following Rollout Start in Phase 1 NY, Phase 2, and Phase 3



Notes: IV estimates of the impact of mailing the Ticket in the rollout month are indicated relative to the average for all rollout months. All estimates are constrained to sum to zero in each phase. The slope of each linear trend line reflects the estimates from the same model with linear restrictions imposed on the coefficients; the coefficients, slope estimates, and test statistics are reported in Appendix C, Exhibit C.4.

For the Phase 1 NY sample (left panel in Exhibit III.4), there is clearly no evidence of an effect of duration to MM on TWP completion as of month 48. For both Phase 2 (center panel) and Phase 3 (right panel), the monthly estimates are jointly significant at the 5 percent level. As we demonstrate below, however, the patterns of monthly coefficients in each phase do not support the conclusion that their joint significance reflects an impact of duration to MM on TWP completion.

For the Phase 2 sample, the monthly estimates are jointly significant primarily because only two estimates have relatively large magnitudes—one positive and one negative. However, the first, for March, is negative and the second, for April, is positive. We would expect the opposite to be true if duration to MM reduces TWP completion. Further the trend line has a very small and insignificant slope. It appears that the significant estimates for these two months are simply due to chance.

For the Phase 3 sample, the null hypothesis of all zero coefficients is rejected again because of two coefficients with relatively large magnitudes—one positive and one negative; the fitted line is steeper than in Phase 2 because the difference in coefficients for these two months is in the expected direction and they are further apart from each other. The slope of the trend line is statistically significant, -0.040 percentage points per month. In interpreting this evidence, however, it is important to recognize that the pattern of monthly estimates for Phase 3 is as irregular as it is for Phase 2. Further, the hypothesis that the impacts are linear is rejected. This leaves open the distinct possibility that the estimates for Phase 3 are simply due to chance rather than to a negative impact of duration to MM on TWP completion.

C. Unclear Evidence of Impacts on STW

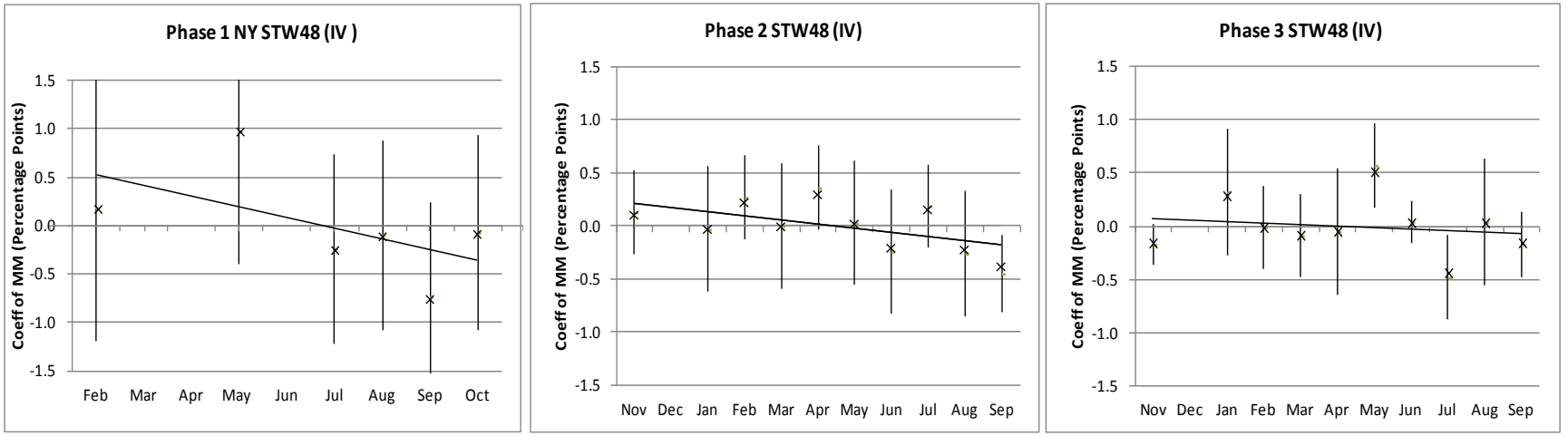
For STW, we found weak evidence that duration to MM reduced attainment of STW as of month 48 for Phase 2, but no such evidence for other phases. Unlike the corresponding impact estimates for service enrollment (Exhibit III.3), the estimates do not display patterns consistent with an impact of duration to MM on STW as of the outcome month. The evidence of an impact on STW as of months 12, 24, and 36, presented in Appendix E, Table E1d, is weaker.

In Exhibit III.5, we present the IV estimates for the impact of duration to MM on STW as of 48 months following rollout start. The linear trend lines have substantial negative slopes for the Phase 1 NY (left panel) and Phase 2 samples (center panel), but the trend line for Phase 3 has only a very small negative slope (right panel). The slope of the line for Phase 1 NY is not statistically significant, despite its substantial size (-0.109 percentage points per month), and we are also unable to reject the hypothesis that all of the monthly estimates are zero.

Viewed in isolation, the Phase 2 results are consistent with the hypothesis that the duration to MM has a substantial negative impact on STW at 48 months. The slope of the line for Phase 2 is less than half as large as for Phase 1 (-0.046 percentage points per month), but it is significant at the 10 percent level. Extrapolation of this estimate implies that a 12-month delay in mailing the Ticket after the rollout start would have reduced STW completion as of month 48 by a substantial amount: 0.6 percentage points. The hypothesis that all coefficients are zero is also rejected at the 10 percent level, and we do not reject the hypothesis that the impact of duration is linear.

The Phase 3 results do not reinforce this interpretation of the Phase 2 findings, however. The slope of the line for Phase 3 is much smaller (0.015 percentage points) and not at all significant. We do find that the monthly Phase 3 estimates are jointly significant, but the insignificant slope of the trend line and the irregular pattern of the monthly estimates suggest that this is due to chance. This leaves open the distinct possibility that Phase 2 estimates also simply reflect chance.

Exhibit III.5. Estimated Impacts of Duration to MM on STW at 48 Months Following Rollout Start in Phase 1 NY, Phase 2, and Phase 3



Notes: IV estimates of the impact of mailing the Ticket in the rollout month are indicated relative to the average for all rollout months. All estimates are constrained to sum to zero in each phase. The slope of each linear trend line reflects the estimates from the same model with linear restrictions imposed on the coefficients; the coefficients, slope estimates, and test statistics are reported in Appendix C, Exhibit C.6.

Although the results for Phase 2 are consistent with a substantial negative effect, they are not statistically strong. The Phase 1 NY and Phase 3 results provide very weak support, at best. The results in Phases 2 and 3 for STW also seem inconsistent with those for TWP completion. As reported earlier, the point estimate for the impact on TWP completion is roughly five times larger in Phase 3 than Phase 2 (based on the slope of the line when the linear restrictions are imposed), but the point estimate for the impact on STW is about three times larger in Phase 2 than Phase 3. In other words, the estimates suggest that for Phase 3, a substantial impact on TWP completion did not translate into a substantial impact on STW, whereas in Phase 2 a much more modest impact on TWP completion translated into a substantial impact on attainment of STW. A possible explanation of these inconsistent results is that they are all due to chance.

D. Unclear Evidence of Impacts on NSTW Months

In Exhibit III.6, we plot the monthly IV estimates for impacts of duration to MM on the number of NSTW months at 48 month after rollout start. As with STW, the NSTW months evidence from Phase 2 is marginally indicative of a substantive impact when viewed in isolation, but in the context of all findings there is a distinct possibility that the Phase 2 results simply reflect chance.

For the Phase 1 NY sample (left panel), there is clearly no evidence of an effect of duration to MM on NSTW months as of month 48. For Phase 2, the estimates are jointly significant at the 5 percent level (center panel). We do not reject the null hypothesis that the impact of duration to MM is linear over the rollout period. However, the slope of the trend line is not very large[18] and is statistically insignificant. For Phase 3 (right panel), we reject the hypothesis that all monthly values are zero at the 1 percent level. However, the pattern of coefficients is indicative of *positive* effects, rather than negative—as illustrated by the positive slope of the trend line—and we reject the hypothesis that the impact of duration to MM is linear over the rollout period at the 5 percent level. Thus, it seems likely that the Phase 3 results simply reflect chance rather than an impact of duration to MM, and chance might equally well explain the Phase 2 results.

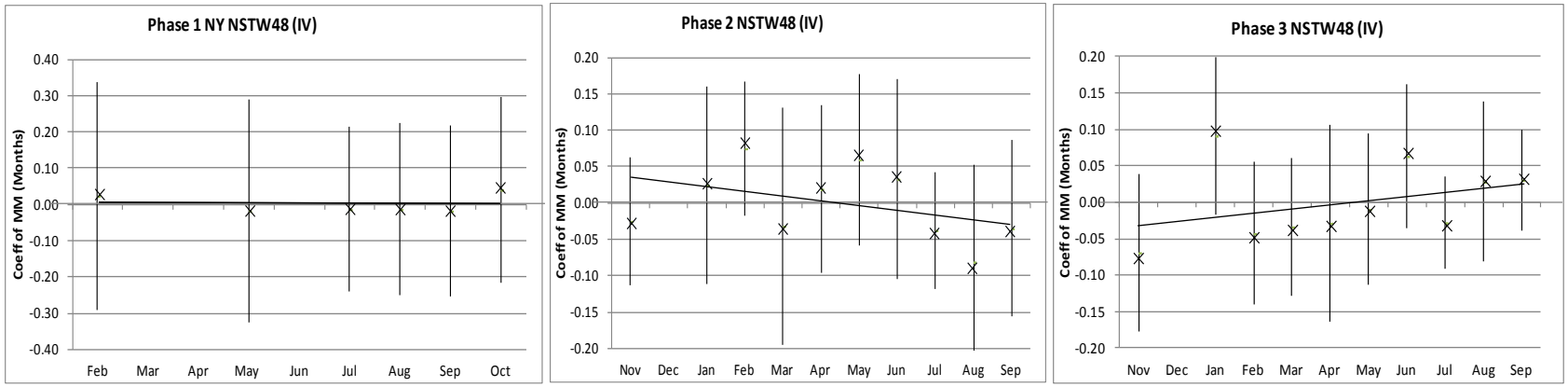
In summary, as with the STW results, the Phase 2 results for NSTW months are consistent with the hypothesis of small negative impacts of duration to MM. The results are not statistically strong, however, and the conclusion that the impact was negative is undermined by the absence of any evidence consistent with negative impacts in the Phase 1 NY and Phase 3 samples as well as the inconsistent evidence of impacts on TWP completion in the Phase 2 and 3 samples. The analysis of total impacts, presented in the next section, reinforces the conclusion that the marginally significant impacts on STW and NSTW months found for Phase 2 are simply the result of chance.

E. Projections of Total Impacts of TTW

Estimates for the impacts of the duration to MM on any outcome at later points in time can be converted to projections of "total impacts" if certain assumptions are maintained. That is, the projections are estimates of cumulative impacts as of that later month of mailing the Ticket in the first rollout month versus not mailing it at all. We call these estimates projections because of their

[18] The slope is -0.0059 NSTW months per month of duration to MM. This translates into an effect of -0.7 months for a 12-month delay in mailing the Ticket—over a 48-month period.

Exhibit III.6. Estimated Impacts of Duration to MM on NSTW Months at 48 Months Following Rollout Start in Phase 1 NY, Phase 2, and Phase 3



Notes: IV estimates of the impact of mailing the Ticket in the rollout month are indicated relative to the average for all rollout months. All estimates are constrained to sum to zero in each phase. The slope of each linear trend line reflects the estimates from the same model with linear restrictions imposed on the coefficients; the coefficients, slope estimates, and test statistics are reported in Appendix C, Exhibit C.8.

reliance on two maintained assumptions. First, the marginal impact of delaying the mailing of the Ticket on service enrollment (or other outcome) as of month 12, 24, 36, or 48 is linear through month 13 of the 48-month observation period for each sample (hereafter the "linearity" assumption). That is, we extrapolate the fitted line to month 13 only, but do so at each observation point.

The second assumption is that the impact of mailing the Ticket on service enrollment for those mailed Tickets in month 13 is always exactly 12 months behind the impact on enrollment for those mailed Tickets in month one (hereafter the "total impact only delayed" assumption). For instance, the impact of mailing Tickets in month 13 as of month 24, 36, or 48 is exactly the same as the impact of mailing the Ticket in month one as of month 12, 24, or 36, respectively. A detailed explanation of how the projections are constructed under these assumptions appears in Appendix D Section A.

Under this approach, there could be some upward bias in the projection, although we think it is likely to be quite small. If the two maintained assumptions (linearity and total impact only delayed) are correct, the projection is unbiased. It seems likely, however, that the second of the two maintained assumptions is somewhat optimistic. Specifically, we would expect impacts for those mailed Tickets in month 13, if anything, to be somewhat smaller than for those mailed Tickets in month one because of the passage of 12 months. During this period, their human capital might well have deteriorated, they might have become better adapted to living on benefits and not engaging in SGA, or they might have managed to find a job or increase their earnings without assistance. If so, then the impact of delay as of 24 months overstates the incremental impact between months 13 and 24 of mailing the Ticket in month one instead of month 13, because part of the impact of the delay is the negative effect of waiting on the size of the total impact for those mailed Tickets in month 13. An analogous statement applies to the estimates for the impact of duration to MM over the other 12-month intervals.

We applied this approach to all outcome variables in all four samples (See Appendix D, Exhibit D.2). The projected total impact for service enrollment as of 48 months for Phase 1 NY is 5.5 percentage points, but it is only significant at the 10 percent level because of a relatively large standard error (SE)—reflecting the relatively short rollout period and relatively small sample. The projected impact for Phase 2 just half as large, 2.3 percentage points, but is statistically significant. The projected impact for Phase 3 is more modest at 1.2 percentage points, but it is also significant.

A final feature of the service enrollment projections is that the point estimates increase with the projection month in each phase—reflecting the maintained assumptions and the fact that the restricted IV estimates of all coefficients in the duration to MM models are positive. Further, for Phases 2 and 3, the increment to the projection diminishes with each 12-month period, as we would expect. That is not true for two Phase 1 samples, likely reflecting the relatively large SEs for those estimates.

None of the projections for total impacts on other outcome variables are significant at even the 10 percent level as of any observation point, with the exception of one marginally significant estimate with a sign that is opposite to what was expected (see Appendix D, Exhibit D.2). To illustrate, consider the projections for NSTW months. We expected these projections to be positive—especially given the impact on service enrollment—but more are negative than positive, and with one exception all are not close to being statistically significant. The one marginally significant projection has a sign opposite that expected, for month 12 in Phase 1 NY (-0.1 month). The NSTW-months estimates stand in stark contrast to those for service enrollment—the latter with

uniformly positive point estimates and, apart from the Phase 1 Except NY projections, significant at the 0.10 level or better. These projections reinforce our earlier conclusion that there is no evidence of a substantial impact on any outcomes other than service enrollment.

F. Assessment of the Hypothesis That TTW Was Self-Financing by 2007

The fact that we did not find statistically significant impacts on STW or NSTW months does not by itself rule out the possibility that TTW under the initial regulations had impacts on these outcomes that were sufficiently large for the program to be "self-financing"—that is, for savings from a net reduction in benefits to be sufficient to pay for TTW payments to providers and all administrative costs attributed to the program. Thornton (2012) suggests that only a very small impact—an increase of 3,000 or so in the number of all beneficiaries experiencing STW for the first time in each year—might be sufficient for the program to be self-financing. An annual impact on first-time STWs that is as small as 3,000 might correspond to such a small impact on STW as of month 48 for new, young SSD beneficiaries that the evaluation would be unable to differentiate between that impact and no impact at all. This section summarizes our assessment of whether the evidence from the above analysis allows us to confidently rule out the possibility that TTW was self-financing in 2007—the last full calendar year prior to the change in the regulations. See Appendix D, Section C for a more detailed discussion of the related issues.

An impact of 3,000 is quite small relative to the number of first-time STW cases actually observed in any recent year. Based on findings in Schimmel et al. (2013) and additional tabulations of their data, we estimate that an impact of 3,000 first STW cases is about five percent of the number of first STW cases in 2007 that would have occurred in the absence of TTW. Under certain strong assumptions, we could conclude that a TTW impact of five percent or greater on STW at 48 months for new, young SSD-only beneficiaries would be large enough to have made the program self-financing in 2007. As discussed in more detail in Appendix D, the most critical of these assumptions are: (1) that rapid growth in the number of beneficiaries during the period leading up to 2007 did not increase the number of STW cases in 2007 relative to the number that would be observed for a stable beneficiary population of the same size; (2) that any impact on delivery of tickets during the rollout to beneficiaries who had been on the rolls for many years had dissipated before 2007; and (3) that the percentage impact of TTW on STW for young SSD-only beneficiaries is essentially the same as for other age-program groups. We consider these assumptions in greater detail in Appendix D, Section C, and conclude that under more realistic assumptions, a larger percentage impact would likely be required for the program to be self-financing.

Because we have no means of knowing precisely how much above five percent the necessary self-financing impact is likely to be under more realistic assumptions, we treat the five percent figure as a lower bound and test the following hypothesis: the mailing of tickets to young, new SSD-only beneficiaries increased the number who had attained STW as of month 48 after the mailing by at least five percent versus the alternative hypothesis that the impact was less than five percent. We repeat the test for NSTW months, on the assumption that an increase in STW of five percent would be sufficient for TTW to be self-financing only if NSTW increases by at least the same relative amount; if those who attain STW as the result of TTW return to the rolls quickly rather than accumulating NSTW months, reductions in benefits would be minimal. Finally, we consider how the results would change if the minimum percentage impact consistent with self-financing was larger than five percent, as it might well be.

We again focus on Phases 2 and 3 because the power of the STW and NSTW projections for these two phases is much greater than for Phase 1; it is clear that the power of the Phase 1

projections is insufficient to rule out an impact of the required size. We consider the Phase 2 and 3 projections separately, and then, to increase power, we pool the results for the two phases on the assumption that the true relative impacts for the two phases are the same. The pooled projection is the minimum variance projection under the assumption that percentage impacts were the same for Phases 2 and 3.[19] Because of the inequalities in the null and alternative hypotheses, a one-tailed test is appropriate. Results appear in Exhibit III.7. We also show tests for the null hypothesis of "no impact" versus the one-tailed alternative of "positive impact."

For STW, the percentage projections for both phases and the pooled percentage projection are all larger than 5 percent (barely so in Phase 3: 5.4 percent). Hence, in each case we are not able to reject the null hypothesis that the impact is 5 percent or larger. For the pooled projection, the p-value for the test is 0.74—far above the 0.10 value that is the usual standard for marginal rejection of the null hypothesis. Note, however, that we are also unable to reject the null hypothesis that the impact is zero versus the alternative that it is positive, although the p-value for this test based on the pooled data is much closer to 0.10: 0.14. In short, for this outcome, the evidence is more consistent with an impact of at least 5 percent than with an impact that is zero or negative.

In contrast, for NSTW months, the percentage projections are all *smaller* than 5 percent, and both the Phase 3 and pooled projections are negative (-8.0 percent and -3.4 percent, respectively). We cannot, however, reject the null hypothesis of a 5 percent impact based on the pooled sample (p-value of 0.14). In this case, however, the p-value is much smaller than the p-value for the test of the null hypothesis that the true impact is zero (0.67 percent). That is, for NSTW months the evidence is more consistent with the hypothesis of a zero or negative impact than with an impact of at least 5 percent.

Exhibit III.7. Projected Relative Impacts on STW and NSTW at 48 Months After Mailing

	Phase 2	Phase 3	Pooled
STW			
Projected relative impacts at 48 months	22.1%	5.4%	11.6%
Standard error of relative impacts	17.8%	13.7%	10.9%
P-value for test of "no impact (or negative impact)" versus "positive impact"	0.108	0.346	0.142
P-value for test of "impact of 5.0% (or more)" versus "impact less than 5.0%"	0.831	0.513	0.729
NSTW			
Projected relative impacts at 48 months	4.1%	-8.0%	-3.4%
Standard error of relative impacts	12.5%	9.7%	7.7%
P-value for test of "no impact (or negative impact)" versus "positive impact"	0.371	0.795	0.673
P-value for test of "impact of 5.0% (or more)" versus "impact less than 5.0%"	0.471	0.090	0.136

Notes: The relative projected impacts were calculated by comparing the projected total impacts on STW and NSTW as of 48 months (see Appendix D, Exhibit D.2) and the estimated means for the corresponding mean in the absence of TTW (counterfactual). For each outcome, the counterfactual mean was estimated by subtracting the weighted mean of the Phase 2 and 3 impact estimates at 48 months from the actual mean for the phase. The p-values are for one-tailed tests, reflecting the inequalities in the hypotheses.

[19] The minimum variance estimate is a weighted mean of the estimates for the two phases where the weights have been chosen to minimize the variance of the estimate. More weight is given to the Phase 3 estimate for each impact because the Phase 3 estimate has lower variance than the Phase 2 estimate.

In summary of the analysis to this point, under the strong assumptions discussed above, the statistical power of the projections for STW and NSTW months is insufficient to rule out the possibility that TTW had impacts of at least five percent on each outcome for Phases 2 and 3 pooled. At the same time, the evidence from these projections alone is just as consistent with zero or negative impacts. The pooled projection for STW of 11.6 percent clearly is more consistent with the hypothesis of an impact of at least 5 percent, but the pooled projection for NSTW months of -3.4 percent is more consistent with a zero or negative impact.

There are reasons to believe that the smallest percentage impact estimated for 2007 that is consistent with self-financing is larger than five percent (see Appendix D, Section C.1). Because of this, a five percent impact represents a lower bound on the impact necessary for TTW to be self-financing. If we had used a larger value in the tests above, the results would be less favorable to the hypothesis of self-financing. The value used would have to be several times larger for the STW test to lead to rejection of the null-hypothesis at the five-percent significance level when using the pooled data: 30 percent. At the same time, however, the value used would only need to be nine percent for the NSTW test to lead to rejection of the hypothesis of self-financing at the same significance level, again using the pooled data.

Thus, although the results overall are consistent with no impact, and we are confident that the estimates are not biased, the statistical power of the methodology is not sufficient to definitively discriminate between "no effect" and "smallest effect consistent with self-financing." If we allow for more realistic assumptions—especially the likelihood that rapid growth in the number of SSD beneficiaries leading up to 2007 substantially inflated the number of first STW cases in that year relative to the number we would expect to observe in a stable beneficiary population of the same size and with the same characteristics—then the impact necessary to reach self-financing would need to be higher. If self-financing required at least a nine percent impact on NSTW months—a plausible value—we would have to reject the hypothesis that TTW was self-financing as of 2007.

IV. CONCLUSION

This final chapter includes a summary of the findings and a discussion of what they add to our understanding of how the TTW affected service enrollment, earnings, and benefit outcomes under the initial regulations.

A. Summary of Findings

We find clear evidence that the mailing of Tickets during the rollout period did increase service enrollment. However, we found very little evidence that the impact translated into a substantive increase in suspension or termination for work (STW) or in the number of NSTW months. Key estimates of impacts of duration to MM on the latter two outcomes from the Phase 2 rollout are marginally significant and consistent with a substantive impact, but findings from other phases and for intermediate outcomes—TWP start and completion—in Phase 2 do not reinforce the Phase 2 findings, and are instead consistent with the interpretation that the Phase 2 findings are simply due to chance. Projections of total impacts at 48 months reinforce this interpretation.

The most important findings pertain to how the amount of time between the start of the rollout and the mail month (MM)—the month in which the SSA mailed a Ticket to the beneficiary—impacts service enrollment in the first 12 months. The Phase 2 and 3 findings are very significant and consistent with each other; the Phase 2 point estimate implies that a 12-month delay in Ticket mailing reduces enrollment by 1.0 percentage point; the Phase 3 estimate implies a reduction of 0.8 of a percentage point. These estimates are substantial. The Phase 2 value is 15 percent of the mean for the percent of those in the sample who enrolled for services over the entire 48-month observation period (6.8 percent); and the Phase 3 value is approximately the same percentage of the corresponding mean for that sample (5.7 percent). The point estimate for the Phase 1 NY sample is much larger—2.5 percentage points—and is nearly 30 percent of the percentage enrolled for services as of month 48 (8.5 percent), but also has a much wider confidence interval due to the much smaller sample and shorter rollout period. We attribute the lack of a significant finding for the Phase 1 Except NY sample to the very short rollout period.

Another important aspect of the findings regarding service enrollment is that little evidence exists of an effect of duration to MM on service enrollment by the end of the 48-month observation period; it appears that those who were mailed Tickets late within each rollout period essentially caught up—in terms of enrolling for services—with those who were mailed Tickets earlier in the rollout period.

It is also notable that the point estimates of impacts on service enrollment decline from the first of the rollout samples to the last. This might be due to chance, but there is an alternative explanation associated with the fact that, on average, the later the rollout sample, the longer a beneficiary had been on the SSD rolls when SSA mailed his or her Ticket. The reason for the decline in the impact estimates across successive phases might be that the longer after SSD entry SSA waits to mail Tickets the lower the impact. That would be consistent with findings that SSD beneficiaries who return to work and complete the TWP are most likely to do so during their first five years on the rolls (Liu and Stapleton 2011).

We also projected the (cumulative) total impacts of mailing the Ticket in the first rollout month on service enrollment as of 12, 24, 36, and 48 months under two maintained assumptions that are quite plausible. We project that the total impact on service enrollment at month 48 is considerably

larger than the impact as of month 12: 5.5 percentage points in Phase 1 Except NY (versus 2.6 as of month 12)): 2.3 percentage points in Phase 2 (versus 1.0) and 1.2 percentage points in Phase 3 (versus 0.8).

The analysis provides no consistent evidence of impacts on other outcomes. Some estimates for Phase 2 are suggestive of an impact, but it seems likely that they are due to chance. Specifically, marginally significant Phase 2 point estimates for STW and NSTW-months imply that a 12-month delay in mailing a Ticket increases attainment of STW as of 48 months by 0.6 percentage points and increases number of NSTW months by an average of 0.07 months. The estimate for STW is about 7 percent of the corresponding percentage for the whole sample at the end of 48 months (8.1 percent), and the estimate for NSTW months is about 5 percent of the mean for NSTW months at the end of 48 months (1.46 months).

Although the Phase 2 estimates on their own are indicative of an impact of Ticket on these outcomes, there are substantial reasons to believe that the results are simply due to chance. The fundamental reason is that whenever an evaluation produces impacts for many outcomes, there are bound to be a few statistically significant findings even if the intervention has absolutely no impacts. That reflects the design of statistical tests: using a 5 percent significance level means that we will have a 5 percent chance of rejecting the null hypothesis of no effect for any individual impact, even if the null hypothesis is true. Hence, if we tested impacts for 100 independent outcomes, we would expect to find significant impacts for five outcomes even if there were no impacts at all. We have produced impact estimates for many different outcomes (not all independent), so we would expect to find that some estimated impacts beyond those for service enrollment would be statistically significant even if there are no impacts on these outcomes. Hence, to assess whether the Phase 2 results for STW and NSTW months reflect real impacts or simply chance, it is important to consider them in the context of all the estimates produced—are the latter consistent with real impacts for these outcomes in Phase 2?[20]

The comparison of the Phase 3 estimates for STW and NSTW months to those for Phase 2 is particularly problematic because of the very large Phase 3 sample and a rollout period that is just as long as that for Phase 2. Think of Phase 3 as an independent test of the same intervention on a population that differs only slightly from the Phase 2 population (the Phase 3 beneficiaries reside in different states and have been on the rolls somewhat longer). The Phase 3 estimates imply an impact on STW at 48 months that is just one-third the size of the estimate for Phase 2, and an impact on NSTW months that is in the opposite direction found for Phase 2 and just as large. We can as easily interpret the Phase 3 estimates as evidence that Ticket mailing reduced NSTW months as we can interpret the Phase 2 estimates to be evidence that Ticket mailing increased NSTW months. It is very hard to understand why comparable impacts on service enrollment in the two samples would translate into such different impacts for NSTW months.

Results for TWP completion also undermine the conclusion that the Phase 2 estimates for STW and NSTW months reflect real impacts. We find almost no relationship between duration to MM and TWP completion in Phase 2, and marginally significant evidence of a negative impact of duration to MM on TWP in Phase 3. This is the opposite of what we would expect if the Phase 2

[20] There are formal ways to address the multiple comparison problem (see Schochet 2008, 2009). We have not conducted a more formal analysis because so few estimates other than those for service enrollment are even marginally significant.

estimates for STW and NSTW months reflect real impacts, and if the Phase 3 estimates for the same variables indicate an absence of real impacts in Phase 3.

Examination of the plots of coefficients for individual MM in Chapter III also reveals why we sometimes find significant evidence of impacts, not necessarily in the expected direction, in the results for a single outcome. Each month's coefficient is a random estimate. It is often the case that one or two of these estimates (10 in the case of Phase 2 and Phase 3) are positive or negative outliers relative to the others, which we would expect to occur by chance. If an estimate for a month early in the rollout period happens to be a positive outlier and one for a later month happens to be a negative outlier, a regression line fit to the point estimates is likely to have a negative slope; that is, it will appear that the relationship between duration to MM and the outcome is negative. That clearly happens for STW and NSTW months in Phase 2. But if the opposite occurs—a negative outlier appears in an early month and a positive outlier in a later month—the relationship appears to be positive, as happens for NSTW months in Phase 3.

Finally, the projections for outcome variables other than service enrollment provide no evidence of positive total impacts for these variables as of any of the four observation months. Most notably, estimates of cumulative impacts of TTW on STW and number of NSTW months for Phases 2 are not significantly different from zero as of any of the four observation points.

The consistency of the findings for service enrollment at 12 months across the three samples with the most variation in duration to MM convinces us that those estimates reflect real impacts. Symmetrically, the inconsistency of the findings on TWP start, TWP completion, STW and NSTW months across these same samples implies that there is no evidence of positive impacts for these outcomes.

Although we did not find positive evidence of impacts on STW and the number of STW months, the estimates are not strong enough to rule out the possibility that TTW was self-financing in 2007. Because a very small positive impact of TTW on STW and the number of NSTW months could be sufficient for TTW to pay for itself through benefit reductions, we cannot reject the self-financing hypothesis on the basis of projected impacts for either STW or NSTW months. Thus, although the results overall are consistent with no impact, and we are confident that the estimates are not biased, the statistical power of the methodology is not sufficient to definitively discriminate between "no effect" and "smallest effect consistent with self-financing."

B. Contribution of the Findings to Knowledge About Ticket Impacts

It is worthwhile to consider the analysis and findings presented here in the larger context of efforts to evaluate the impact of TTW.

The Ticket Act directed SSA to implement TTW nationwide without any prior testing. National implementation posed a major challenge to the evaluation, as it limited the options for comparison groups. Initially SSA considered the option of using within-phase variation in MMs to estimate impacts, as we have done here, but at the time the option was unattractive because: (1) SSA had planned short rollout periods for each phase, like that in Phase 1 Except NY; (2) the administrative data did not include monthly measures of a key outcome—months without benefits following suspension or termination for work; and (3) another seemingly attractive option was available.

An additional challenge to the impact evaluation was that any credible impact evaluation required an intent-to-treat approach: comparison of outcomes for all beneficiaries mailed Tickets—

not just those who used their Tickets—to outcomes for a comparison group, because there was no credible approach to identifying those in the comparison group who would have used their Ticket had they received one. The expectation was that only 5 percent of the beneficiaries mailed Tickets would actually assign them. Mean impacts measured over all beneficiaries were bound to be very small because there would be no impact for a large majority of beneficiaries.

The TTW evaluation team pursued the more promising alternative evaluation design: comparison of outcomes in early-rollout states to contemporaneous outcomes in later-rollout states during the rollout period (Thornton et al. 2007; Stapleton et al. 2008). The evaluation used administrative data for millions of beneficiaries, rather than a sample, to maximize the ability to detect small impacts. It focused on annual impacts, rather than monthly impacts, because the best available earnings data were the annual data in the Master Earnings File.

Like the results presented here, the earlier impact analysis found significant impacts on service enrollment. In fact, for SSD-only beneficiaries under age 40, the point estimates were quite comparable to those found here: a 0.6 percentage point increase in service enrollment by the end of the rollout year and a 1.5 percentage point increase at the end of the following year (Thornton et al. 2007; Stapleton et al. 2008). There was also some concern that these estimates were biased upward, however, because of incomplete data on service enrollment; nonetheless, the evidence was quite convincing that there was at least a small impact on service enrollment.[21] The service enrollment estimates presented here substantially increase our confidence that there were impacts on service enrollment, and there is little room for doubt about the size.

The earlier analysis also found what at first appeared to be evidence of impacts on earnings and benefits, but this evidence was undermined by methodological limitations. Specifically, when we applied the same methodology to beneficiaries in the same states during the pre-TTW period (a Heckman-Hotz test), we found similar impact estimates where, of course, no impacts had occurred.[22] We concluded that the impact estimates for the rollout period might well reflect differential trends in mean employment and benefits across the three TTW rollout phases that existed prior to the rollout. Another substantial limitation of the original methodology for the benefit and earnings impact estimates is that it limited the analysis to outcomes observed in the year the ticket was mailed and the following year, while the expectation was that it might take a longer time for impacts on service enrollment to be translated into impacts on earnings and benefits. Hence, the results were ambiguous; we were neither convinced that there were impacts on these outcomes, nor convinced that any impacts were so small as to be inconsequential for policy purposes.

The ambiguous results for earnings and benefits led to a re-examination of options for providing more definitive evidence of impacts under the original regulations. The lengthening of the rollout periods in NY and, especially, the Phase 2 and 3 states; the development of the monthly STW and NSTW months variables; and the passage of enough time to extend the observation period for every phase to 48 months made the use of random within-phase variation in IMM to estimate initial impacts more appealing than when SSA first considered it.

[21] Estimated impacts on service enrollment were significant for young beneficiaries in other program groups, as well, but not as large as for young SSD-only beneficiaries. Estimates for older beneficiaries were smaller and less significant.

[22] See Heckman and Hotz (1989).

As with the new results for service enrollment, the results for other outcomes increase our knowledge about impacts on earnings and benefits. The new analysis does not consider those outcomes directly, but impacts on earnings that are large enough to be of policy interest would be large enough to substantively increase STW and NSTW months, otherwise TTW would not generate substantial benefit savings.

Although we have considerable confidence that the methodology provides unbiased estimates and can detect impacts as small as those found for service enrollment, we have also illustrated that its power is not sufficient to rule out the small positive impacts on STW and NSTW months that would be required for TTW to have been self-financing in 2007. While application of the methodology did not yield positive evidence that the impacts for these variables were that large, it also did not rule out the possibility that they were.

It is important to keep in mind that these estimates are for TTW under the original regulations. The 2008 changes to the TTW regulations clearly stimulated provider interest and the number of beneficiaries assigning their Tickets. In principle, those changes could have had a positive impact on STW and NSTW months. However, it appears impossible to rigorously measure any such impact because the regulations were implemented nationally, without a test, and also because implementation of the new regulations occurred during a recession that was deeper and longer lasting than any recession that has occurred since the 1956 inception of DI.

Finally, the analysis provides a lesson for SSA and other national or state agencies when, in the future, they are asked to make a significant change to a large program at a national or state level—including significant future changes to TTW. Inasmuch as such a change often requires a lengthy rollout period, the agency should consider the knowledge that might be gained by implementing a rollout in which program participants are randomly assigned an implementation month over a period of 12 months or so. Such a randomized rollout might be very attractive if the knowledge to be gained is substantial, and it is otherwise practical, as it was for TTW. This approach has its limits, however; it will not necessarily have sufficient power to identify substantively important impacts if such impacts are very small. The power of the approach can be enhanced if the program participants most likely to be affected by the change can be identified in advance, the rollout period can be lengthened, or a more extreme version of the change could be applied to randomly chosen participants.[23] Such enhancements make this approach more like the approach that would be best from a purely methodological perspective: a randomized control trial in which those likely to be affected by the change are randomly assigned to either a treatment group or a control group, and mean outcomes for the two groups are compared over a period of sufficient length to estimate impacts of interest to the agency.

[23] As an illustration of how an extreme version of a change could be tested, SSA could have initially provided "super tickets" to a randomly chosen group of beneficiaries—for instance, tickets with milestone and outcome payments that are 50 percent larger than those for other tickets.

This page has been left blank for double-sided copying.

REFERENCES

Angrist, J., G. Imbens, and D. Rubin. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association*, vol. 91, no. 434, 1996, pp. 444–455.

Bell, Stephen, and M. C. Bradley. "Calculating Long-Run Impacts of Social Programs with Staggered Implementation." *Empirical Economics*, Online First™, 23 August 2012.

Heckman, James, and Joseph Hotz. "Choosing Among Alternative Nonexperimental Methods for Estimating the Impact of Social Programs: The Case of Manpower Training." *Journal of the American Statistical Association*, vol. 84, 1989.

Hemmeter, Jeff, Jackie Kauff, and David Wittenburg. "Changing Circumstances: Experiences of Child SSI Recipients Before and After Their Age-18 Redetermination for Adult Benefits." *Journal of Vocational Rehabilitation*, vol. 30, no. 3, 2009, pp. 201–221.

Liu, Su, and David Stapleton. "Longitudinal Statistics on Work Activity and Use of Employment Supports for New Social Security Disability Insurance Beneficiaries." *Social Security Bulletin*, vol. 71, no. 3, August 2011, pp. 11–75.

Mamun, Arif, Paul O'Leary, David Wittenburg, and Jesse Gregory. "Employment Among Social Security Disability Program Beneficiaries: 1996-2007." *Social Security Bulletin*, vol. 71, no. 3, 2011, pp. 11–34.

Social Security Administration. *Annual Statistical Report on the Social Security Disability Insurance Program, 2011*. Washington, DC, July 2012.

Schimmel, Jody, David Stapleton, David Mann, Sarah Prenovitz, and Dawn Phelps. "Participant and Provider Outcomes since the Inception of Ticket to Work and an Assessment of the Effect of the Changes to the Regulations in 2008." Draft report submitted to the Social Security Administration. Washington, DC: Mathematica Policy Research, June 2013.

Schochet, Peter Z. "Guidelines for Multiple Testing in Impact Evaluations of Educational Interventions." Final report submitted to the U.S. Department of Education, Institute of Education Sciences. Princeton, NJ: Mathematical Policy Research, May 2008.

Schochet, Peter Z. "An Approach for Addressing the Multiple Testing Problem in Social Policy Impact Evaluations." *Evaluation Review,* vol. 33, no. 6, December 2009.

Stapleton, David, and Arif Mamun. "Early Impacts of Ticket to Work for Young New DI Awardees." Memorandum submitted to Social Security Administration; Washington, DC: Mathematica Policy Research, July 18, 2012.

Stapleton, David, Gina Livermore, Bonnie O'Day, Craig Thornton, Robert Weathers, Krista Harrison, So O'Neil, Emily Sama Martin, David Wittenburg, and Debra Wright. "Ticket to Work at the Crossroads: A Solid Foundation with an Uncertain Future." Report submitted to the Social Security Administration. Washington, DC: Mathematica Policy Research, September 2008.

Stapleton, David, Gina Livermore, Mark Nowak, Ellen Bouchery, and John Kregel. "Evaluation Design for the Ticket to Work Program." Falls Church, VA: The Lewin Group, June 2002.

Stata. "Obtaining Robust Variance Estimates" in *Stata User's Guide: Release 12*. College Station, TXL StataCorp, LP.

Thornton, Craig, Gina Livermore, David Stapleton, John Kregel, Tim Silva, Bonnie O'Day, Thomas Fraker, W. Grant Revell, Jr., Heather Schroeder, and Meredith Edwards. "Evaluation of the Ticket to Work Program Initial Evaluation Report." Washington, DC: Mathematica Policy Research, February 2004.

Thornton, Craig, Robert Weathers, David Wittenburg, Thomas Fraker, David Stapleton, Jesse Gregory, and Arif Mamun. "Initial Impacts of the Ticket to Work Program on Social Security Disability Beneficiary Service Enrollment, Earnings and Benefits." *Journal of Vocational Rehabilitation*, vol. 27, no. 2, 2007, pp. 129–140.

Thornton, Craig, Gina Livermore, Thomas Fraker, David Stapleton, Bonnie O'Day, David Wittenburg, Robert Weathers, Nanette Goodman, Tim Silva, Emily Sama Martin, Jesse Gregory, Debra Wright, and Arif Mamun. "Evaluation of the Ticket to Work Program: Assessment of Post-Rollout Implementation and Early Impacts." Washington, DC: Mathematica Policy Research, May 2007.

Thornton, Craig. "Can the Ticket to Work Program Be Self-Financing?" Washington, DC: Mathematica Policy Research, 2012.

Whalen, Denise, Gina Livermore, and David Stapleton. "Impact Analysis of the Ticket to Work Program Using the NBS Interview Attempt as an Instrument for Ticket Assignment." Memorandum submitted to Social Security Administration, Washington, DC: Mathematica Policy Research, September 23, 2011.

Wooldridge, Jeffrey. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: The MIT Press, 2002.

APPENDIX A

DESCRIPTION OF DATA

This page has been left blank for double-sided copying.

In this appendix, we describe the data sources and define the key variables used in the study. We also present the criteria for inclusion in our analytical samples, describe the sample, and present results from tests of the statistical equivalence of the intended mail month samples within each phase. We also assess variation in the business cycle over the sample period and present the unemployment rate measures that we use in our analysis to assess the effect of the economy on Ticket impacts. Finally, we present descriptive statistics on the outcome measures.

A. Ticket Research File

We used data from the 2007 TRF (TRF07). The TRF is a set of analytic administrative data files constructed for the TTW evaluation. The TRF07 files contain current and historical information on more than 22 million SSD or SSI beneficiaries who received a benefit in at least one month from January 1996 through December 2007 (Hildebrand et al. 2009).[24] For the purpose of this study, we constructed annual cohort files for those awarded benefits from 1999 through 2003.[25] Cohort assignment is based on the month that SSA first paid a benefit to the awardee. Although it is possible for an individual to have multiple entitlements, he or she is assigned to just one cohort based on the year that corresponds to the individual's *first* payment.[26] All analyses were conducted using pooled data from multiple cohorts.

B. Analytic Sample

1. Sample Selection

The sample of interest includes beneficiaries who entered the SSD rolls from July 1999 through October 2003. For the analysis, we followed each beneficiary for 48 months starting with the first month of the rollout in the beneficiary's state. As the Phase 3 rollout started in November 2003, the last month in the sample is October 2007. We started with July 1999 SSD awardees because this is the month in which the nonblind SGA level was increased from $500 to $700. We end the follow-up period in 2007 because of the severe recession that started in the last quarter of 2007 and because SSA made substantial changes to the TTW regulations in 2008, which may have affected beneficiary outcomes in 2008 and later.

The analysis samples consist of young (ages 18 to 39 at award) SSD-only awardees who were first paid SSD benefits no earlier than July 1999 and were selected for the initial rollout of the TTW program on one of three Ticket selection dates: January 12, 2002 (Phase 1), October 26, 2002 (Phase 2), or October 18, 2003 (Phase 3).[27] It was SSA's intent to mail Tickets to every beneficiary

[24] Extracts from several Social Security administrative files were merged to create the TRF, including the Master Beneficiary Record, Supplemental Security Record, Numerical Identification System (Numident) file, the 831 and 832/33 Disability files, the Disability Control File, monthly snapshot files, and files from the payment history update system.

[25] These annual cohort files are an extension of those created by Liu and Stapleton (2011).

[26] The first payment month (that is, the award month) is the month in which the first payment was actually made, which is usually after the first month for which the beneficiary is entitled to a benefit (that is, the entitlement month). The latter is often used in SSA's statistics to classify beneficiaries by entry year (for example, SSA 2009). We use the award month instead because our focus is on the activities of beneficiaries once they become informed of their award and are entitled to use the DI work incentives.

[27] SSA determined all beneficiaries who were eligible to receive a Ticket and who resided within the phase's states during as of the phase's selection month. Almost all SSD and SSI disability beneficiaries over age 18 were eligible; the main significant exceptions were 1) new beneficiaries with a status of Medical Improvement Expected (MIE) who had *(continued)*

in these samples during a subsequent rollout month (hereafter, the "intended mail month" [IMM]), to be determined by the terminal digit of the beneficiary's SSN. As will be seen, SSA mailed the vast majority of these Tickets on the IMM. For each phase, we treat the samples defined by the IMM (hereafter, the "IMM samples") as randomly assigned samples of those included on the phase's selection date.

The number of SSD and SSI beneficiaries included in each selection date sample appears in the first column of Exhibit A.1. From these, we included only beneficiaries who were recent young SSD awardees in July 1999 or later (column 2), and among the latter we excluded those who were also SSI beneficiaries at the time of award (column 3) to obtain the analytic samples (column 4).

Exhibit A.1. Ticket Selection Sample Sizes by Phase

Phase	Ticket Selection Dates	Number of SSD and SSI Beneficiaries Selected	Number of New Young SSD Beneficiaries	Less Concurrent Beneficiaries	Number in Analytic Sample
1	Jan 12, 2002	2,375,970	104,760	49,657	55,103
2	Oct 26 2002	2,577,672	142,379	65,218	77,161
3	Oct 18 2003	3,415,502	211,031	96,374	114,657
Total		**8,369,144**	**458,170**	**211,249**	**246,921**

Note: New young SSD beneficiaries are those who received their first payment in July 1999 or later and who were under age 40 in that month.

Using SSA's internal documentation, we were able to specify the IMM value for each beneficiary in our sample, depending on the terminal digit of their SSN and their Ticket selection date (that is, the Ticket rollout phase). Because a vast majority of the beneficiaries were mailed a Ticket on the month SSA intended, we were also able to cross check the correspondence between a terminal digit and the IMM for each of the phase samples using the actual mail date.

The sample sizes by IMM are displayed in Exhibit A.2. We separated the Phase 1 samples into NY residents and residents from other Phase 1 states because the rollout proceeded on a different schedule in NY than for other states. The months are numbered in three ways: rollout month, starting with 1 for the first month of the phase's rollout; intended mail month; and calendar month.

In each phase, there was a one-month pause in mailings after the first mail month. For administrative reasons, the duration of the Phase 1 rollout was different in NY and other Phase 1 states. For NY, 6 mailings were spread over 9 months; for the rest of Phase 1, 4 mailings were spread over 5 months. Because beneficiary interest overwhelmed providers and SSA's implementation contractor early in the Phase 1 rollout, SSA compensated by reducing the speed in Phases 2 and 3. Phases 2 and 3 follow identical 11-month schedules, except separated by 12 months. Both designs called for uniform distribution of the mailings across 10 of the 11 rollout months, with the second month being the exception. Our samples reflect that design.

(continued)
not yet had their first medical continuing disability review (medical CDR), and 2) SSI child beneficiaries who had reached 18 and were waiting for redetermination as adults.

Exhibit A.2. IMM Sample Sizes by Phase

	Phase 1 NY			Phase 1 Except NY			Phase 2			Phase 3		
Rollout Month	Intended Mail Month	Calendar Month	N	Intended Mail Month	Calendar Month	N	Intended Mail Month	Calendar Month	N	Intended Mail Month	Calendar Month	N
1	1	Feb-02	1,148	1	Feb-02	4,282	10	Nov-02	7,573	21	Nov-03	11,531
2												
3				3	Apr-02	8,644	12	Jan-03	7,733	23	Jan-04	11,328
4	4	May-02	1,163	4	May-02	12,960	13	Feb-03	7,679	24	Feb-04	11,539
5				5	Jun-02	17,194	14	Mar-03	7,745	25	Mar-04	11,569
6	6	Jul-02	2,438				15	Apr-03	7,743	26	Apr-04	11,523
7	7	Aug-02	2,408				16	May-03	7,778	27	May-04	11,533
8	8	Sep-02	2,452				17	Jun-03	7,598	28	Jun-04	11,434
9	9	Oct-02	2,414				18	Jul-03	7,826	29	Jul-04	11,519
10							19	Aug-03	7,681	30	Aug-04	11,253
11							20	Sep-03	7,805	31	Sep-04	11,428
Total			**12,023**			**43,080**			**77,161**			**114,657**

Although SSA actually mailed Tickets on the IMM for most of the beneficiaries, for a small fraction the actual mail month (MM) did not correspond to the IMM (see Exhibit A.3). The TRF records include data on the actual Ticket mail dates, and for each beneficiary we used the first recorded mail date to identify the actual month the Ticket was mailed. Across the four phase samples, 93 to 99 percent of the cases in our samples were mailed a Ticket on the IMM. It is reassuring to see that no case was mailed a Ticket before the rollout started in that phase. However, a small fraction of beneficiaries were first mailed a Ticket before their IMM. This could have happened if a beneficiary had become aware of the TTW program, requested a Ticket at a local SSA office, and received a Ticket prior to their IMM (reflecting a provision of the regulations called "Ticket on demand"). A very small fraction of beneficiaries was mailed a Ticket after the IMM, but before rollout was completed in their state. Another very small fraction of beneficiaries was mailed a Ticket after rollout was completed in their state, which could happen when individuals moved to a state where Ticket was rolled out later. For up to 1 percent of the beneficiaries in the four phase samples, the actual Ticket mail date is missing, indicating that a Ticket was not mailed to them. After examining these cases further, we found that in most cases the beneficiary was deceased as of the IMM; in most other cases the beneficiary was not in current pay status for reasons other than work.

Although the fraction of Tickets mailed on the IMM was very high in each month of the rollout, it did decline in successive months (Exhibit A.4). This happened primarily because of "Ticket on demand," as beneficiaries assigned to a later IMM had longer time to request a Ticket on their own. In addition, it appears that, as the rollout progressed, SSA identified some beneficiaries who had died or were no longer in current-pay status, and consequently did not mail Tickets. Because mailing a Ticket on demand, mortality, and loss of current pay status for some other reason are likely predictive of the beneficiary outcomes of interest, these factors pose challenges for estimating impacts of duration to Ticket mailing on beneficiary outcomes. We address this challenge in our analysis using instrumental variables approach described in greater detail in Appendix B.

2. Beneficiary Characteristics and Tests of Statistical Equivalence

In Exhibit A.5, we present characteristics of the beneficiaries in the four phase samples, treating the NY and Phase 1 except NY samples as separate samples. These variables are used as control variables in the models described in Appendix B. The exhibit also presents tests of the statistical equivalence of the IMM samples within each phase—tests that each phase's selection sample was allocated to the IMM samples for the phase in a manner that was equivalent to random assignment (apart from the deterministic separation of NY cases from other cases in the Phase 1 selection sample).

Almost all of the characteristics are defined as of the beneficiary's Ticket selection date. The exceptions are the primary disabling conditions, measured at SSD award date; the primary insurance amount, which is the earliest recorded value; and the indexed monthly earnings, also the earliest recorded value. The beneficiary populations vary somewhat across phases, as reflected in modest differences in means. Differences between Phase 2 and 3 are of special interest because of their comparable rollouts and size. Compared to the Phase 2 sample, the Phase 3 sample has relatively fewer African Americans (16 percent versus 21 percent), more Hispanics (12 percent versus 3 percent), higher indexed monthly earnings ($1,125 versus $1,090) and Primary Insurance Amount ($643 versus $632), and more beneficiaries with major affective disorders (18 percent versus 16 percent). Many differences reflect the fact that the Phase 3 rollout started 12 months after the Phase 2 rollout, so beneficiaries in Phase 3 had aged a year between the Phase 2 Ticket selection date and their own selection date, and more new awardees were added to the Phase 3 sample during the same

Exhibit A.3. Correspondence of Actual Mail Months (MM) and Intended Mail Months (IMM)

Actual Mail Month Is—	Phase 1 NY		Phase 1 Except NY		Phase 2		Phase 3	
Before Rollout	0	0.00%	0	0.00%	0	0.00%	0	0.00%
Before IMM, During Rollout	248	2.06%	131	0.30%	3,355	4.35%	6,307	5.50%
Corresponds to the IMM	11,661	96.99%	42,549	98.77%	73,008	94.62%	106,938	93.27%
After IMM, During Rollout	24	0.20%	177	0.41%	224	0.29%	39	0.03%
After Rollout	7	0.06%	19	0.04%	51	0.07%	173	0.19%
Missing Mail Date								
Status as of IMM	83	0.69%	204	0.47%	523	0.68%	1,200	1.06%
Deceased	79	0.66%	193	0.45%	433	0.56%	973	0.85%
Suspense/termination for work	0	0.00%	2	<0.01%	2	<0.01%	3	<0.01%
Other	**3**	0.02%	5	0.01%	85	0.11%	204	0.18%
Suspense/termination								
Current pay	1	0.01%	4	0.01%	3	<0.01%	20	0.02%
Total	12,023	100.00%	43,080	100.00%	77,161	100.00%	114,657	100.0%

Exhibit A.4. Percentage Mailed a Ticket on the Intended Mail Month by IMM and Phase

	Phase 1 NY		Phase 1 Except NY		Phase 2		Phase 3	
Rollout Month	IMM	% Actually Mailed on IMM	IMM	% Actually Mailed on IMM	IMM	% Actually Mailed on IMM	IMM	% Actually Mailed on IMM
1	1	99.74	1	99.95	10	97.46	21	99.85
2								
3			3	98.91	12	95.93	23	94.13
4	4	98.11	4	98.69	13	95.86	24	93.89
5			5	98.46	14	95.36	25	93.22
6	6	96.96			15	94.56	26	93.04
7	7	96.84			16	94.33	27	92.39
8	8	96.45			17	94.04	28	92.21
9	9	95.86			18	93.05	29	91.67
10					19	92.98	30	91.42
11					20	92.68	31	90.79

Exhibit A.5. Beneficiary Characteristics: Means by Mail Month for the RV Sample by Phase

	Phase 1 NY			Phase 1 Except NY			Phase 2			Phase 3		
	Mean	Range[a]	Test[b]	Mean	Range[a]	Test[b]	Mean	Range[a]	Test[b]	Mean	Range[a]	Test[b]
Gender												
Male	0.522	0.021		0.521	0.018	***	0.521	0.020		0.523	0.015	***
Gender missing	0.000	0.000		0.000	0.000		0.000	0.000		0.000	0.000	
Age at Selection Date	33.51	0.73		32.89	0.13		33.697	0.192	***	34.31	0.21	**
Age at Disability Onset												
Age	27.70	0.36		27.48	0.17		27.963	0.343	**	27.96	0.29	***
Age missing	0.000	0.000		0.000	0.000	*	0.000	0.000		0.000	0.000	
Race/Ethnicity												
Asian	0.018	0.009		0.013	0.002		0.009	0.005	***	0.026	0.004	
African American (non-Hispanic)	0.186	0.028		0.163	0.010	*	0.209	0.012		0.156	0.011	***
Hispanic	0.093	0.024		0.072	0.007	***	0.032	0.006		0.119	0.011	***
American Indian/Hawaiian	0.003	0.007		0.009	0.002	**	0.006	0.002		0.005	0.002	**
White [Ref]	0.658	0.071		0.721	0.006		0.728	0.016	**	0.667	0.013	**
Other	0.014	0.010		0.007	0.001		0.004	0.002		0.010	0.003	***
Missing	0.028	0.021	**	0.014	0.002		0.012	0.004		0.017	0.004	
Education at Selection Date												
Less than high school [Ref]	0.148	0.040		0.181	0.013		0.187	0.016	***	0.152	0.014	***
High school graduate	0.348	0.072		0.408	0.008		0.391	0.015		0.374	0.013	***
More than high school	0.227	0.034		0.208	0.007		0.185	0.018	***	0.204	0.012	***
Missing	0.278	0.100	*	0.203	0.012	***	0.238	0.013		0.270	0.014	***
Expectations about Medical Improvement at Selection Date												
Expected	0.033	0.009		0.030	0.005	**	0.039	0.009		0.028	0.004	
Possible	0.511	0.093	*	0.580	0.005		0.548	0.015	***	0.601	0.014	***
Not expected	0.150	0.051		0.189	0.007		0.180	0.016	***	0.196	0.011	*
Missing	0.306	0.117		0.201	0.011	**	0.233	0.013		0.175	0.014	***
Medicare Eligibility at Selection Date [Ref = not eligible]												
Eligible	0.579	0.089		0.500	0.016		0.646	0.011		0.779	0.015	***
Eligibility missing	0.035	0.009		0.035	0.004		0.031	0.006		0.018	0.004	***
Number of Dependent Beneficiaries at Selection Date												
No dependent [Ref]	0.459	0.054		0.499	0.012		0.455	0.017	***	0.478	0.017	**
1	0.145	0.028		0.150	0.006		0.171	0.011	**	0.163	0.010	***
2 or more	0.255	0.033		0.233	0.007		0.265	0.012		0.249	0.013	***
Missing	0.141	0.036		0.118	0.006		0.108	0.012	**	0.110	0.007	***
VR Services Before Selection Date												
Determined eligible for VR services	0.186	0.012		0.196	0.006		0.191	0.018	***	0.207	0.010	
VR service eligibility missing	0.742	0.010		0.738	0.003		0.761	0.024	***	0.760	0.011	

Exhibit A.5 *(continued)*

	Phase 1 NY			Phase 1 Except NY			Phase 2			Phase 3		
	Mean	Range[a]	Test[b]	Mean	Range[a]	Test[b]	Mean	Range[a]	Test[b]	Mean	Range[a]	Test[b]
Outcome Achieved Before Ticket Selection Date												
TWP start before Ticket selection	0.052316	0.019106	*	0.061	0.003		0.063	0.010	***	0.084	0.010	**
TWP completion before Ticket selection	0.030774	0.012289		0.040	0.005		0.038	0.007	***	0.048	0.007	**
STW before Ticket selection	0.016468	0.005761		0.017	0.004	**	0.018	0.004		0.025	0.005	**
Months Between SSD Award and Selection Date	13.625	1.013		12.974	0.233	*	18.496	0.224		26.352	0.540	***
Primary Insurance Amount (PIA, $)												
Mean PIA	646.7	60.4		595.0	10.8		626.9	18.8	*	643.3	12.3	***
PIA missing	0.157	0.031		0.139	0.004		0.125	0.011	**	0.123	0.009	***
Indexed Monthly Earnings (IME, $)												
Mean IME	1180.1	205.0		1023.6	23.3		1089.6	45.8		1124.9	28.4	***
IME missing	0.245	0.051		0.205	0.006		0.192	0.009		0.185	0.008	**
Primary Disabling Conditions at SSD Award												
Major affective disorders [Ref]	0.114	0.055		0.160	0.009	**	0.155	0.018	***	0.184	0.015	
Other psychiatric disorders and mental retardation	0.234	0.052		0.272	0.012		0.241	0.019	**	0.241	0.017	***
Back disorders and musculoskeletal system	0.199	0.110		0.100	0.005		0.120	0.009	**	0.111	0.011	***
Other physical disabilities	0.452	0.036		0.468	0.011		0.482	0.027	***	0.463	0.017	***
Missing	0.001	0.002		0.001	0.001	**	0.001	0.001		0.001	0.001	
SSD Award Year												
1999	0.153	0.047		0.125	0.007		0.097	0.005		0.077	0.009	***
2000	0.372	0.026		0.371	0.003		0.328	0.012		0.258	0.009	
2001	0.373	0.028		0.395	0.012		0.300	0.012	*	0.293	0.014	
2002	0.076	0.018		0.085	0.005		0.250	0.013		0.232	0.016	***
2003	0.014	0.007		0.014	0.005	*	0.015	0.004	***	0.128	0.015	***
2004	0.013	0.007		0.011	0.003		0.009	0.003	*	0.012	0.003	***
State Unemployment Rate												
Mean in 6 months around IMM	0.071	0.700		-0.030	0.253	***	0.079	0.843	***	-0.255	0.264	***
Change in 6 months around IMM	6.158	0.510		5.740	0.032	***	5.710	0.385	***	6.007	0.511	***
Phase 1 States												
Arizona	0.000	0.085		0.085	0.004		0.002	0.002		0.001	0.001	
Colorado	0.000	0.042		0.042	0.003		0.001	0.001		0.001	0.001	
Delaware	0.000	0.017		0.017	0.003		0.001	0.000		0.000	0.000	
Florida	0.000	0.243		0.243	0.009		0.007	0.003		0.004	0.002	
Iowa	0.000	0.044		0.044	0.005		0.001	0.001		0.001	0.001	
Illinois	0.000	0.182		0.182	0.003		0.004	0.003		0.002	0.001	

Exhibit A.5 *(continued)*

	Phase 1 NY			Phase 1 Except NY			Phase 2			Phase 3		
	Mean	Range[a]	Test[b]	Mean	Range[a]	Test[b]	Mean	Range[a]	Test[b]	Mean	Range[a]	Test[b]
Massachusetts	0.000	0.112		0.112	0.004		0.005	0.002		0.003	0.001	
New York	1.000	1.000		0.000	0.000		0.009	0.002		0.005	0.003	
Oklahoma	0.000	0.055		0.055	0.006		0.001	0.002		0.000	0.001	
Oregon	0.000	0.042		0.042	0.006		0.001	0.001		0.001	0.001	
South Carolina	0.000	0.071		0.071	0.012		0.003	0.002		0.002	0.002	
Vermont	0.000	0.010		0.010	0.001		0.000	0.001		0.000	0.000	
Wisconsin	0.000	0.076		0.076	0.008		0.002	0.002		0.001	0.001	
Phase 2 States												
Alaska	0.000	0.000		0.000	0.000		0.006	0.003		0.000	0.000	
Arkansas	0.000	0.000		0.000	0.000		0.040	0.006		0.001	0.001	
Connecticut	0.000	0.001		0.001	0.001		0.038	0.004		0.001	0.002	
District of Columbia	0.000	0.000		0.000	0.000		0.005	0.004		0.000	0.000	
Georgia	0.000	0.001		0.001	0.000		0.089	0.011		0.002	0.001	
Indiana	0.000	0.001		0.001	0.001		0.075	0.008		0.001	0.001	
Kansas	0.000	0.000		0.000	0.001		0.025	0.004		0.001	0.001	
Kentucky	0.000	0.000		0.000	0.000		0.070	0.013		0.002	0.001	
Louisiana	0.000	0.000		0.000	0.000		0.051	0.006		0.001	0.001	
Michigan	0.000	0.001		0.001	0.001		0.123	0.007		0.004	0.002	
Missouri	0.000	0.001		0.001	0.001		0.072	0.008		0.003	0.002	
Mississippi	0.000	0.000		0.000	0.000		0.047	0.005		0.002	0.002	
Montana	0.000	0.000		0.000	0.000		0.007	0.002		0.000	0.000	
North Dakota	0.000	0.000		0.000	0.000		0.005	0.002		0.000	0.000	
New Hampshire	0.000	0.000		0.000	0.000		0.019	0.004		0.001	0.001	
New Jersey	0.000	0.001		0.001	0.001		0.081	0.015		0.003	0.001	
New Mexico	0.000	0.000		0.000	0.000		0.017	0.004		0.000	0.001	
Nevada	0.000	0.001		0.001	0.001		0.020	0.005		0.001	0.001	
South Dakota	0.000	0.000		0.000	0.000		0.006	0.002		0.000	0.000	
Tennessee	0.000	0.001		0.001	0.000		0.077	0.009		0.002	0.002	
Virginia	0.000	0.001		0.001	0.001		0.080	0.010		0.002	0.001	
Phase 3 States												
Alabama	0.000	0.000		0.000	0.001		0.001	0.001		0.051	0.006	
California	0.000	0.002		0.002	0.001		0.001	0.001		0.190	0.012	
Hawaii	0.000	0.000		0.000	0.000		0.000	0.000		0.008	0.002	
Idaho	0.000	0.000		0.000	0.000		0.000	0.001		0.010	0.004	
Maryland	0.000	0.000		0.000	0.000		0.001	0.001		0.036	0.005	
Maine	0.000	0.000		0.000	0.000		0.000	0.001		0.018	0.003	
Minnesota	0.000	0.001		0.001	0.000		0.000	0.001		0.041	0.003	
North Carolina	0.000	0.001		0.001	0.001		0.001	0.001		0.082	0.010	
Nebraska	0.000	0.000		0.000	0.000		0.000	0.000		0.012	0.003	
Ohio	0.000	0.001		0.001	0.001		0.001	0.001		0.092	0.008	
Pennsylvania	0.000	0.001		0.001	0.001		0.001	0.001		0.123	0.007	
Puerto Rico	0.000	0.000		0.000	0.000		0.000	0.001		0.063	0.005	

Exhibit A.5 *(continued)*

	Phase 1 NY			Phase 1 Except NY			Phase 2			Phase 3		
	Mean	Range[a]	Test[b]	Mean	Range[a]	Test[b]	Mean	Range[a]	Test[b]	Mean	Range[a]	Test[b]
Rhode Island	0.000	0.000		0.000	0.000		0.000	0.000		0.011	0.002	
Texas	0.000	0.001		0.001	0.001		0.001	0.001		0.130	0.008	
Utah	0.000	0.000		0.000	0.000		0.000	0.000		0.012	0.002	
Washington	0.000	0.001		0.001	0.001		0.000	0.001		0.045	0.007	
West Virginia	0.000	0.000		0.000	0.000		0.000	0.001		0.024	0.005	
Wyoming	0.000	0.000		0.000	0.000		0.000	0.000		0.003	0.001	
U.S. Territories	0.000	0.000		0.000	0.000		0.000	0.000		0.002	0.001	
Anomalous Sequence of Events												
Ticket selection before SSD award	0.075	0.016		0.080	0.014	***	0.036	0.008	**	-	-	
TWP completion before start	0.0001	0.0004		0.000	0.000	**	0.000	0.001		0.000	0.000	***
STW before TWP start	0.0002	0.0004		0.000	0.000		0.000	0.000		0.000	0.000	***
STW before TWP completion	0.0013	0.0008		0.002	0.001	***	0.001	0.002	***	0.001	0.001	***

Source: Authors' calculations using data from Ticket Research File 2007

Note: "Ref" indicates the reference category for the discrete variable in the multivariate regression models.

[a]"Range" is the difference between the minimum and maximum mean across IMM in each sample.
[b]"Test" shows the results from the test of the null hypothesis that the means are the same across IMM in the sample; *, **, and *** indicate significantly different at the 0.10, 0.05, and 0.01 levels, respectively.

period. Compared to those in the Phase 2 sample, as of the selection date, they were older (mean of 34.3 versus 33.7) and had been on the rolls longer (mean of 26 months versus 18 months). In addition, those in the Phase 2 sample were more likely to have: previously enrolled for services, started the TWP, completed the TWP, experienced a month of suspension or termination for work, and become eligible for Medicare.[28] In addition, some differences are expected between Phases 2 and 3 because of differences between the economic, policy and cultural environments for states in each phase.

The statistical equivalence tests for each phase's sample were conducted by running linear regressions of each characteristic on a set of IMM indicators for the months within that phase, without an intercept. Thus, for instance, for Phase 1 we include indicators for February, April, May, and June of 2002. For each regression, we conducted a joint test (an F-test) for the hypothesis that all of the population coefficients are equal. In conducting the test, we treated each state in the phase as a cluster and allowed for heteroscedasticity in the regression disturbance.

Sample means by phase along with the p-values for the corresponding F-test within the phase appear in Exhibit A.5. The F-tests show that we would reject the null hypothesis of "no difference" across IMM samples within phase for a large number of characteristics, especially in the largest samples: Phases 2 and 3. Substantively, however, even when a baseline characteristic is found to be statistically different across IMM samples within a phase, variation in the means across the IMMs is not substantial. For example, in Phase 2, we found significant differences in mean beneficiary age at Ticket selection date across IMMs, but the difference between the maximum and minimum mean is 0.19 years around a mean for the phase of 33.70. Perhaps most bothersome, there is significant variation in the achievement of some of the event outcomes across IMM within phase that is substantial relative to the overall mean. For instance, in Phase 2, 19.1 percent of beneficiaries had previously been found eligible for SVRA services, and the range of this percentage across the IMM was 1.8 percent.

We consider the IMM samples in each phase to be substantively very similar even though the joint tests indicate that they are not statistically equivalent in many regards. The differences likely reflect the fact that the method SSA used to assign IMM was not purely random, particularly with respect to age and factors that are associated with age. For this reason, it is important to control for these characteristics in the analysis—most critically, for the occurrence of the outcome events prior to Ticket selection date. We did this by including dummy variables for the occurrence of the outcome events prior to Ticket selection date in the estimated model.

In each phase, a small fraction of the sample resides in states that were not targeted for rollout in that phase. This is because we defined Ticket rollout phase based on the Ticket selection date (January 12, 2002 is Phase 1, October 26, 2002 is Phase 2, and October 18, 2003 is Phase 3), and did not rely on the beneficiary's state of residence (with the exception of NY). As shown in Exhibit A.6, across the three phase samples with multiple states, 2.2 percent to 4.7 percent of the sample reside in a state which did not correspond with the states targeted for that phase. We do not know the cause. Perhaps SSA included people in neighboring states that were served by a field office located in a state within the phase group; alternatively, the state shown in the data reflects an address

[28] Somewhat disconcerting, in a small share of cases, it appears that SSA made the first benefit payment to the beneficiary after the Ticket selection date; for instance, for the Phase 1 samples, selected late in 2002, a small share of beneficiaries were first paid in 2003 and 2004, according to records. We suspect this reflects errors in the administrative data that led to misidentification of the first payment date.

that is not the beneficiary's own, or there were retroactive changes. Because we are aiming to retain as much of the original IMM sample as possible for the analysis, and because we found little variation in the percentage of the sample in each of these states across the mail months within phase, we did not exclude these cases.[29]

Exhibit A.6. Beneficiary Residence in Phase

	P1 NY		P1 Except NY		P2		P3	
	N	%	N	%	N	%	N	%
Resides in state included in phase	0	0	42,154	97.85	73,546	95.31	109,239	95.27
Resides in state not included in phase	12,023	100	926	2.15	3,615	4.69	5,418	4.73
Total	**12,023**	**100**	**43,080**	**100**	**77,161**	**100**	**114,657**	**100**

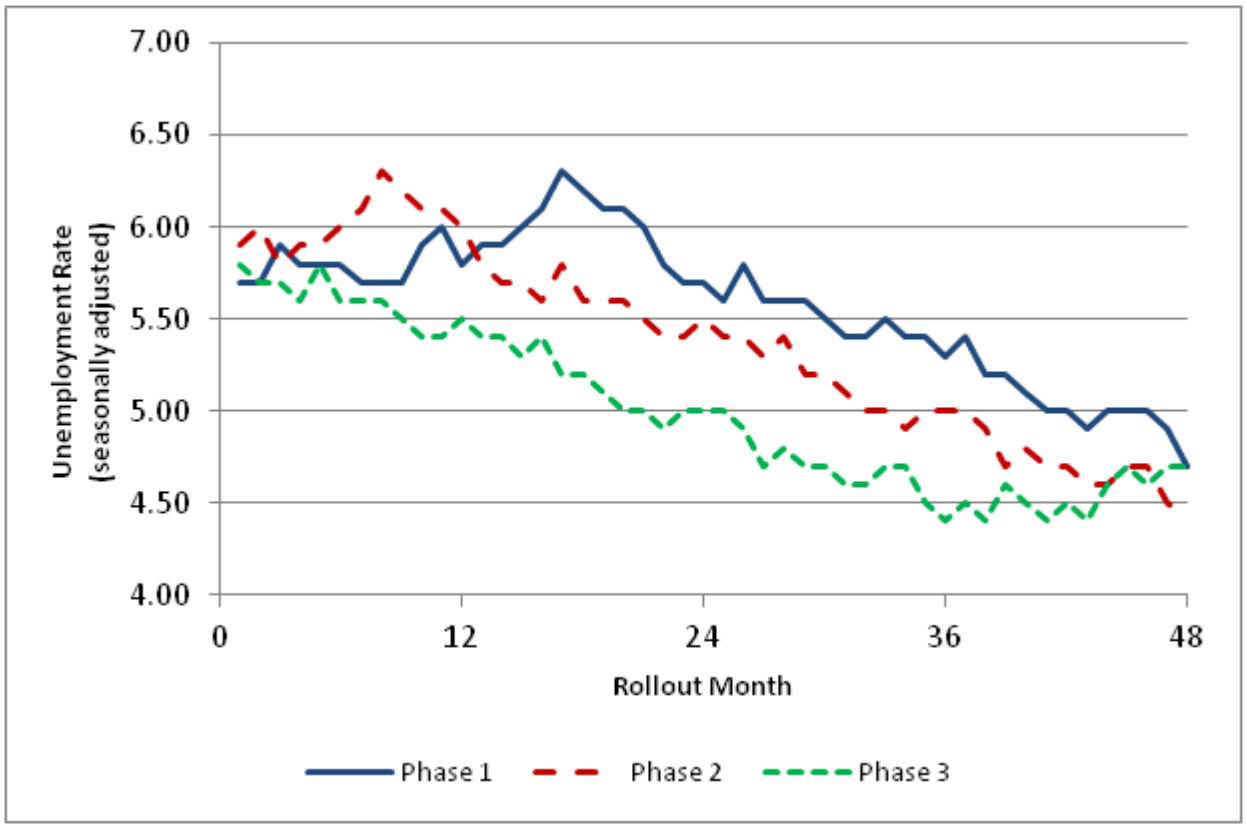
3. The Economy

The Ticket rollout occurred during the economic expansion following the 2001 recession, but the strength and timing of the recovery varied across states and rollout phases. This is illustrated in Exhibit A.7, where we plot monthly national unemployment rates from the first rollout month in each phase through the 48th month (the analysis period for each rollout). In the exhibit we show that all three analysis periods are predominantly periods of recovery, but the timing of recovery varies substantially across phases. The unemployment rate continued to rise in 2002 and did not peak until after the completion of the Phase 1 rollout, in June 2003—approximately the middle of the Phase 2 rollout. The unemployment rate started to decline before the Phase 3 rollout began and continued to decline for the next 36 months, before leveling out and then starting to rise in the last few months of the period as the next recession approached (recall that the 48th month for Phase 3 is October 2007).

Of course, all beneficiaries within each phase were subject to the same national economic circumstances during the analysis period for their phase, so there is no reason to be concerned about bias in the impact estimates from within-phase correlation between a beneficiary's mail month and the economic conditions experienced by the beneficiary. It is at least arguable, however, that the impact of Ticket mailing interacts with the state of the economy at the time the Ticket is mailed. For instance, beneficiaries might be more successful finding work on their own, without a Ticket, during a strong expansion than they would be during the trough of the business cycle.

[29] Another issue with including these cases in the sample comes from the fact that we treat each state as a cluster, and correct the standard error of our estimates for clustering. Considering the small samples in the "out-of-phase" states, each of which is treated as a cluster, it is conceivable that they might be substantially influencing the results of the joint tests of statistical equivalence. To explore this, we conducted the joint tests without correcting for state-level clustering (but adjusting for heteroscedasticity) and found that we would reject the null hypothesis for far fewer characteristics. We suspect that with a small number of cases in the nontargeted states in each phase, the random variation at the cluster level (that is, in the cluster component of the model's error term) explains so much variation in some characteristics that tiny differences across IMM groups are found to be significant. But this is just conjecture, and we are not aware of any technical problem with including a set of clusters with very small samples along with clusters that are much larger.

Exhibit A.7. Unemployment Rate by Rollout Month



To account for variation in economic conditions at the state level around the beneficiary's IMM, we included two aggregate measures of state-level monthly unemployment rates: first, the average unemployment rate during the six-month period around each beneficiary's IMM (from two months before through three months after the IMM), and, second, the change in state monthly unemployment rate during the same period. The means for these measures are shown in Exhibit A.5 above. It appears that beneficiaries in two of the four phase samples—Phase 1 NY and Phase 2—faced a slightly deteriorating economy, on average, when their Tickets were mailed, and those in the other two samples faced a slightly improving economy. We also found significant variation in these variables across the IMM within each phase, indicating that the economic conditions faced by a beneficiary in the beneficiary's IMM did vary across IMM with phase.

4. Outcome Measures

The outcome measures are all based on the 48 months starting with the first rollout month for the phase (month zero is the pre-rollout month). This period ends in January 2006 for Phase 1, September 2006 for Phase 2, and September 2007 for Phase 3. For each individual in the sample we constructed:

- Four binary "event" variables. We determined when in the 48 months following start of rollout each of the following events occurred, if at all: (1) enrolled for employment services (assigned their Ticket to an EN or were determined eligible for services by an SVRA); (2) completed their first TWP month; (3) completed their last TWP month; and (4) had their benefits suspended or terminated for work. In the analysis of whether an event has occurred as of a specified rollout month (month 12, 24, 36, or 48), we define a

binary for each event that is equal to one if the event occurred after the rollout start and before that month, and zero otherwise.

- NSTW months, a count of the number of months in nonpayment status following STW that occurred during the 48-month period. NSTW months include all months after benefits are suspended or terminated for work until the first of the following events occurs: (1) return to current-pay status, (2) suspension or termination for some other reason, or (3) the end of the 48-month period. Beneficiaries are not necessarily engaged in SGA during all NSTW months; we know only that they are not receiving benefits.

We present means for the event variables in the IMM samples as of month 48 in Exhibit A.8 and means for NSTW months as of month 48 by phase in Exhibit A.9. The overall mean for Phase 1 is lower than each of the corresponding means for the two Phase 2 samples, but the mean for Phase 3 is higher than all of the others. The latter difference might be because, compared to the other samples, more Phase 3 beneficiaries had achieved STW prior to the rollout, reflecting the longer time they had been on the rolls prior to the rollout (see Exhibit A.5).

Exhibit A.8. IMM Sample Percentages Experiencing Four Events by End of Month 48 After Rollout Start

	Service Enrollment		TWP Started		TWP Completed		STW	
	%	SE	%	SE	%	SE	%	SE
Phase 1 NY			**11.70**	**0.29**	**9.37**	**0.27**	**8.13**	**0.25**
Feb-02	8.71	0.83	10.98	0.92	8.62	0.83	7.93	0.80
May-02	8.86	0.83	11.44	0.93	10.06	0.88	9.03	0.84
Jul-02	8.45	0.56	11.32	0.64	8.82	0.57	7.79	0.54
Aug-02	8.51	0.57	11.46	0.65	9.18	0.59	7.93	0.55
Sep-02	8.36	0.56	12.93	0.68	9.50	0.59	7.83	0.54
Oct-02	8.49	0.57	11.56	0.65	10.02	0.61	8.66	0.57
Phase 1 Except NY	**8.53**	**0.13**	**11.46**	**0.15**	**9.81**	**0.14**	**8.02**	**0.13**
Feb-02	8.15	0.42	10.81	0.47	9.62	0.45	8.10	0.42
Apr-02	8.82	0.30	11.59	0.34	9.76	0.32	7.81	0.29
May-02	8.74	0.25	11.57	0.28	9.93	0.26	8.17	0.24
Jun-02	8.32	0.21	11.47	0.24	9.80	0.23	7.99	0.21
Phase 2	**6.78**	**0.09**	**9.47**	**0.11**	**8.08**	**0.10**	**6.21**	**0.09**
Nov-02	7.25	0.30	9.45	0.34	7.88	0.31	6.14	0.28
Jan-03	6.91	0.29	9.61	0.34	8.33	0.31	6.21	0.27
Feb-03	6.72	0.29	9.87	0.34	7.96	0.31	6.56	0.28
Mar-03	6.75	0.29	9.37	0.33	7.64	0.30	6.13	0.27
Apr-03	6.70	0.28	9.61	0.33	8.69	0.32	6.53	0.28
May-03	6.78	0.28	9.49	0.33	8.15	0.31	6.31	0.28
Jun-03	6.69	0.29	8.96	0.33	8.08	0.31	6.09	0.27
Jul-03	6.45	0.28	9.47	0.33	8.01	0.31	6.48	0.28
Aug-03	6.46	0.28	9.45	0.33	8.06	0.31	5.87	0.27
Sep-03	7.07	0.29	9.38	0.33	7.99	0.31	5.80	0.26
Phase 3	**5.67**	**0.07**	**8.80**	**0.08**	**8.24**	**0.08**	**6.78**	**0.07**
Nov-03	5.90	0.22	9.20	0.27	8.33	0.26	6.71	0.23
Jan-04	5.73	0.22	9.09	0.27	8.76	0.27	7.07	0.24
Feb-04	5.84	0.22	8.78	0.26	8.22	0.26	6.66	0.23
Mar-04	5.67	0.22	8.56	0.26	8.17	0.25	6.64	0.23
Apr-04	5.39	0.21	8.34	0.26	8.17	0.26	6.78	0.23
May-04	5.54	0.21	9.10	0.27	8.45	0.26	7.40	0.24
Jun-04	6.09	0.22	8.94	0.27	8.26	0.26	6.87	0.24
Jul-04	5.55	0.21	8.67	0.26	7.78	0.25	6.27	0.23
Aug-04	5.46	0.21	8.79	0.27	8.17	0.26	6.78	0.24
Sep-04	5.57	0.21	8.58	0.26	8.12	0.26	6.63	0.23

Note: "%" is the mean percent experiencing the event before the end of the 48 months following the pre-rollout month; SE is the standard error of the estimated mean.

Exhibit A.9. Mean NSTW Months at 48 Months by Phase and Intended Mail Month for the RV Samples

	Phase 1 NY		Phase 1 Except NY			Phase 2		Phase 3	
	Mean	SE	Mean	SE		Mean	SE	Mean	SE
Feb	1.881	0.215	1.834	0.111	Nov	1.412	0.074	1.645	0.065
Mar					Dec				
Apr			1.885	0.080	Jan	1.470	0.075	1.750	0.068
May	1.825	0.203	1.811	0.063	Feb	1.597	0.079	1.617	0.064
Jun			1.814	0.055	Mar	1.394	0.072	1.641	0.065
Jul	1.790	0.143			Apr	1.483	0.075	1.668	0.066
Aug	1.754	0.144			May	1.472	0.074	1.756	0.067
Sep	1.839	0.147			Jun	1.542	0.079	1.733	0.068
Oct	2.087	0.161			Jul	1.448	0.074	1.675	0.067
					Aug	1.357	0.072	1.672	0.066
					Sep	1.423	0.075	1.703	0.067
All Months	1.864	0.067	1.830	0.035		1.460	0.024	1.686	0.021

Notes: All Phase 1 months are in 2002, Phase 2 months start in 2002 and end in 2003, and Phase 3 months start in 2003 and end in 2004. Mean is the mean number of NSTW months over the first 48 rollout months, SE is standard error of the mean.

APPENDIX B

DISCUSSION OF METHODS

This page has been left blank for double-sided copying.

We used two econometric approaches to conduct the impact analysis. We describe our identification strategy in Section A. Then, in Section B, we describe the of use of linear regression models to estimate the effect of duration from rollout start to the *intended* mail month (IMM) on each of the outcome variables, ignoring the fact that not all Tickets were actually mailed in the intended month. In Section C, we discuss the instrumental variables (IV) analogs of the linear regression models to estimate the impacts of duration from the rollout start to the *actual* mail month (MM) on the same outcomes using the IMM as an instrument. We close the appendix with a brief synopsis of other methods that we applied to a preliminary MM data set and an explanation of why we elected to proceed with the methods we chose once the final data set was constructed. The purposes of the last section are to show that the analyses reported here were influenced by previous analysis of a preliminary data set, but only in a manner designed to avoid pre-test bias.

A. Identification Strategy

For each phase of the TTW rollout, SSA selected the intended month of Ticket mailing for all eligible beneficiaries on a Ticket selection date that occurred approximately one month before rollout began and used the terminal digit of the beneficiary's SSN to determine the rollout month in which SSA would mail the beneficiary a Ticket. Because the last four digits (the serial numbers) of SSNs are considered to be random,[30] this strategy essentially led to a random assignment of the eligible beneficiaries to different IMMs, and consequently generated exogenous random variation in the duration from the rollout start to the IMM. This provides the foundation for estimating the impacts of the duration from Ticket rollout start to the IMM on later beneficiary outcomes.

We used the random assignment of beneficiaries to IMMs for identifying two sets of impact estimates. First, we estimate the direct, intent-to-treat (ITT) impacts of delaying the intended Ticket mail month on beneficiary outcomes. We hypothesize that the longer the duration from rollout start to the IMM the less likely the beneficiaries are to achieve the four outcomes—enrollment in vocational services, beginning of TWP, completion of TWP, and attainment of STW—and subsequently the number of months they remain on STW. To estimate the ITT impacts we use linear regression models, with the IMMs on the right hand side along with other beneficiary characteristics prior to Ticket mailing as control variables. We estimated a parallel set of models in which we also included the two unemployment rate variables described in Appendix A. Estimates from the two sets differ little, and standard errors are slightly smaller when the unemployment variables are omitted. We focus on the results without unemployment because both sets of estimates are unbiased and those without unemployment are more efficient.

Second, we estimated the impacts on beneficiary outcomes of delaying the *actual* mailing of Ticket during the rollout period. In a specific sense, these estimates are treatment-on-the-treated (TOT) impacts if we consider the treatment to be the actual mailing of a Ticket. If, instead, we consider the treatment to be *use* of the Ticket by the beneficiary (that is, assignment to a provider in exchange for services), these estimates must also be considered to be ITT estimates, reflecting SSA's randomized initial intent to mail along with intended, nonrandom revisions that occurred subsequent to the Ticket selection date because of Ticket on demand, or loss of current pay status for various other reasons.[31] Methodologically, the revisions to the mail months are like crossover

[30] As noted earlier, the serial numbers in SSNs are considered random only after conditioning on age because they are historically assigned in sequence.

[31] As stated in Chapter II, in order to address the fact that some Tickets were never mailed because benefits were suspended or terminated prior to the IMM, we coded the MM for those observations as if the Tickets were actually *(continued)*

effects—some subjects were offered a treatment that was different from the intended treatment, or were not offered a treatment at all. The first set of estimates is unbiased ITT estimates for the mean impacts of duration to IMM, given the nonrandom revisions that were made as the mailings proceeded. However, they do not answer a more policy-relevant question concerning the impacts of the duration to the *actual* mail month. In fact, the first set of analysis would provide biased estimates of what impacts would have been had all Tickets actually been mailed as originally intended. That bias might be especially severe because we are relying on random variation in duration from rollout start to the mail month to identify impacts, and the later a beneficiary's IMM, the greater the likelihood of an adjustment to the actual MM. We address this potential bias by essentially replacing the IMMs in each model with the actual MMs, then using the IMMs as instrumental variables for the MMs.

The linear regression models and the IV models are discussed in further detail below.

B. Linear Regression Models

To measure the impact of the IMM on duration to each event we estimated linear probability models for the likelihood that the event has occurred by the end of periods of fixed length following the rollout start.[32] We estimated linear probability models for each of the four event variables (service enrollment, beginning TWP, TWP completion, and STW) as of 12, 24, 36, and 48 months after the rollout start, separately for each phase. We also used linear models to analyze the impact of IMM on the cumulated count of NSTW months as of 12, 24, 36, and 48 months after the rollout start, separately for each phase.

Each model is of the following form:

Equation (1) $$E_{it} = \alpha_t + \beta_t' IMM_i + \gamma_t' X_i + \varepsilon_{it}$$

where: E_{it} is a dummy variable for whether the event has occurred for beneficiary i as of month t following the rollout start, or is the number of NSTW months experienced by beneficiary i as of month t following the rollout start; IMM_i is a vector of dummies for the IMM, X_i is the usual vector of baseline control variables, and ε_{it} is a random disturbance. We used the Regression command in Stata to estimate the linear probability models.

When multiple dummy variables like those represented by IMM_i are used in regression models, the usual practice for avoiding exact collinearity is to omit one category (the "base" category) so that each coefficient represents the mean difference between the outcome for the included category and the base category. The choice of base is arbitrary, but in this context could lead to misinterpretation

(continued)
mailed on the IMM. We had previously verified that, with almost no exceptions, termination or suspension of benefits had occurred for reasons other than work—most commonly mortality. Our reasoning for this modification is that mailing the Ticket to these beneficiaries during any month of the rollout would almost certainly have had no impact on their employment outcomes, in which case essentially all outcomes would have been the same as those observed had SSA mailed all of these Tickets in their IMM.

[32] Given the dichotomous nature of the dependent variables, one could use a nonlinear estimation model, such as logistic regression. We used the linear probability model because it provides a convenient approximation of outcome probability around the mean values of the covariates (see, for instance, Wooldridge 2002). The linear probability model is easier to estimate, especially for samples as large as ours, and also has the advantage of a direct and more intuitive interpretation.

of the results if the mean of the outcome variable for the chosen base mail month, by chance, happens to be exceptionally high or low. To avoid that possibility, we instead chose a normalization that is less frequently used, but that is symmetric with respect to the mail months (that is, it does not require that we arbitrarily treat one differently than the others). Specifically, we include dummy variables for all the mail months and impose the restriction that the coefficients of the mail months sum to zero. As a result, the coefficient for each IMM indicates the impact of that MM relative to a typical/average MM during the rollout period.

We assumed that the variance of the disturbance has a component that varies by state as well as a component that varies by individual. We estimated each equation using adjustments for clustering at the state level.[33] To test the null hypothesis that duration to IMM has no impact on an outcome, we performed an F-test for the hypothesis that all of the mail-month coefficients are zero.[34]

For each outcome, we examined whether the coefficients of the mail months show patterns that are consistent with the hypothesis that the longer the duration to mailing the Ticket the less likely the event occurred in the relevant time period, or the fewer the number of NSTW months. We also tested the hypothesis that the marginal impact of delaying the mailing of the Ticket an additional month was the same throughout the rollout period (that is, that there is a linear relationship between duration to mail month and the expected outcome). For this test, each restriction is of the form:

Equation (2) $\beta_{jt} - \beta_{j-1,t} = K,$

where j is an index for rollout month and K is a constant across rollout months. If there is a one-month gap between rollout months, the restriction becomes $\beta_{jt} - \beta_{j-2,t} = 2K$. If there are m rollout months, there are $m - 1$ marginal effects, but only $m - 2$ are independent in the unrestricted model because of the adding-up restriction. Moreover, we are able to test one less restrictions on the $m - 2$ marginal effects. Hence, the number of degrees of freedom for the numerator of the F-test for these restrictions is $m - 3$.

C. Instrumental Variables Estimation

To estimate the impact of actually mailing the Ticket on each outcome, we essentially replace the IMM dummies with MM dummies and use the IMM dummies as instruments. Because the deviations in the MM from the IMM are not random and their number is related to duration from the rollout start to the IMM (as explained above), we used the IV approach to estimate unbiased impacts of mailing the Ticket on each outcomes.

[33] More specifically, we used the "vce (cluster)" option in Stata's regression command, which applies a clustered sandwich estimator and produces standard errors allowing for intracluster correlation, relaxing the usual requirement that the observations be independent (Stata 2012). For the Phase 1 NY sample, we used the "vce (robust)" option in Stata's regression command, which applies the Huber/White/sandwich estimator and produces heteroskedasticity corrected standard errors.

[34] In models where with m dummy variables in IMM_t the number of degrees of freedom (d.f.) in the numerator for this test is $m - 1$, because of the adding-up restriction. The d.f. in the denominator is the number of states. For NY, where there is only one cluster, the d.f. in the denominator is the sample size minus the number of unrestricted coefficients in the model.

Specifically, we estimate models of the form:

Equation (3) $E_{it} = \alpha_t + \beta_t' MM_i + \gamma_t' X_i + u_{it}$

where MM_i is a vector of dummies for the *actual* mail month, all other variables are defined as previously, and u_{it} is the residual. Technically, the bias in the least squares estimates for this model is due to correlation between the dummy variables in MM_i and the equation's unobserved disturbance (that is, v_i). To address this issue, we use the IV approach, which can be represented by the following model:

Equation (4) $MM_i = \lambda_i + \theta\, IMM_i + \tau X_i + v_i$

Equation (5) $E_{it} = \alpha_t + \beta_t' MM_i + \gamma_t' X_i + u_{it}$

where v_i is the residual term, all variables are as previously defined. Because MM_i is a vector, Equation (4) represents a set of equations, one for each mail month, λ_i is a vector of the same length as MM_i, and θ and τ are both matrices.

The specification must address two additional issues. First, some Tickets were mailed in months that were not rollout months. Although we could include additional dummies for these months in the model, these dummies would also have to be treated as endogenous, but then the specification would fail a basic requirement for identification in IV estimation: in the absence of other information, the number of instruments must be at least as large as the number of endogenous variables. To address this problem, we recoded the dummy variable for the last mail month in each phase to equal one if the Ticket was mailed after that month. That is, we have restricted the impact of mailing the Ticket after the last mail month to equal the impact of mailing it in the last mail month. This restriction is inconsequential because much less than 1 percent of Tickets were mailed after the last mail month in each phase, and most of those were mailed within a few months after the last mail month.

Second, some Tickets were not mailed at all. For these cases, all of the mail month dummies were originally coded to be zero. As a result, however, the dummies are no longer exactly collinear—those not mailed a Ticket have become a base category. This is problematic, however. Under this specification, the adding-up condition (coefficients of the mail months sum to zero) is no longer a normalization, but rather a restriction: the mean impact of Ticket mailing across all the mail months in the phase's rollout period is required to be zero relative to not being mailed a Ticket at all.[35]

To address this problem, we redefined the MM dummy values for those never mailed a Ticket to reflect what the values would be if SSA had actually mailed them a Ticket in the initially intended mail month. Thus, for instance, if the July IMM dummy equals one for such a beneficiary, and all others are zero, the July MM dummy is set equal to one and all other MM dummies are set equal to zero. Underlying this specification is an assumption that each outcome for these beneficiaries would

[35] The normalization could be dropped, but then we would be using those not mailed a Ticket, most of whom had died, as the base group; that is, we would be attempting to estimate the impact of mailing a Ticket during the mail month versus death or the other reasons that could lead SSA to not mail the Ticket. Further, the number of instruments would be insufficient because the instruments themselves are exactly collinear (they sum to unity for each observation)—if the number of MM dummies is m the number of independent IMM dummies is $m - 1$.

have been the same if SSA had mailed the Ticket in the month intended, or, for that matter, in any other month. That assumption seems quite reasonable for those who died—the majority of these cases—and also seems plausible for those whose benefits were suspended or terminated for a variety of other reasons.[36]

Two assumptions must be satisfied for IMM_i to be a set of valid instruments.[37] First, conditional on X_i, IMM_i must be uncorrelated with unobserved individual characteristics (including environmental circumstances) or any other factors that affect the outcome variable apart from the actual mail month. Second, again conditional on X_i; IMM_i must be correlated with the MM_i.[38]

Both assumptions are satisfied in our case. The first assumption is satisfied because SSA assigned IMM in a fashion that was in essence random with respect to the individual's characteristics after conditioning on age (one of the variables in X_i); thus, by design, IMM_i is independent of any individual characteristics. Further, the IMM selected could have no effect on the outcomes of interest except through its effect on the actual mailing of Tickets. The second assumption is satisfied because the vast majority of Tickets were mailed on the IMM (as shown in Exhibit II.3). Taken together, the randomly assigned IMMs constitute valid instruments for estimating the impact of timing of actual Ticket mailing on beneficiary outcomes. Further, they are a very strong set of instruments, in that the correlation between the IMM and MM variable for each mail month is very high because SSA mailed the vast majority of the Tickets for each intended mail month in the intended mail month (see Appendix A, Exhibit A.3).

Analogous to the tests conducted for the duration to IMM estimates, for the IV models we conducted tests of the hypothesis that duration to MM has no impact (that is, that the MM coefficients are all zero), as well as tests of the hypothesis that the marginal impact of a one month delay in *actually* mailing the Ticket is the same throughout the rollout period. We used a chi-square test in each case, with d.f. equal to the number of restrictions (the same is the d.f. in the numerator for the corresponding F-test in the linear model estimates for duration to IMM).

D. Methods Applied in Preliminary Analysis

We conducted a number of analyses before we were able to identify the Ticket selection date samples used in the analysis reported here. We initially approximated the samples by identifying everybody in current pay in the month prior to the phase's rollout (January 2002 for Phase 1, October 2002 for Phase 3, and October 2003 for Phase 3) who lived in one of the respective phase's states when they received their first SSD payment, and who were also included in SSA's batch

[36] To provide another perspective, suppose we instead had specified random coefficients for the MM dummies using the original definition of MM. This specification would recognize explicitly the fact that the impact of mailing in each month varies across beneficiaries; it is likely zero for most beneficiaries, but positive and varying for others. When SSA decided to not mail Tickets to some beneficiaries in a month, it primarily removed beneficiaries for whom the effect of mailing would very likely have been zero. As it took time for SSA to identify these cases, many more were removed in the later mail months than in the early mail months—changing the distribution of mail-month effects for the later months by more than for the earlier mail months and increasing via selection the mean effect for those sent a Ticket month. Our redefinition of the MM dummies reverses this effect, so that differences in the mean coefficients across mail months will not reflect the fact that SSA chose withhold mailings to those known to have died or to have lost eligibility for other reasons.

[37] See Angrist et al. 1996.

[38] Technically, the assumptions are (1) IMM_i is uncorrelated with the disturbance terms in equations 4 and 5, and (2) the covariance between MM_i and IMM_i must be of rank $m - 1$.

mailing at the beginning of one the months in the phase's rollout. These sample selection criteria meant that some with randomly assigned IMM were not included in the sample—those SSA dropped from the mailing list because of mortality; or suspension, or termination for other reasons; those mailed a Ticket on demand (normally mailed outside of the batch mailing); and those mailed Tickets in a month outside the rollout period. To partially address the selection problem resulting from Tickets not mailed in the IMM, we excluded beneficiaries who had been mailed a Ticket but were deceased as of the last mail month in the rollout period.

Using this sample, we estimated two types of survival analysis models separately by phase: Kaplan-Meier (nonparametric) survival curves, and Cox proportional hazard models. For NSTW months, considering the nature of the outcome measure, we estimated Poisson (count) regression models. For the purposes of increasing estimator precision, the proportional hazard models, the linear probability models, and the Poisson models included most of the previously mentioned control variables.

In the analysis discussed in this report, we dropped the survival analysis and count-data models, and instead focus on the models discussed above because the latter are less restrictive, the findings are easier to interpret, and they facilitate the use of instrumental variables for purposes of estimating the effects of the actual mail month rather than the IMM.

The preliminary analyses yielded: consistently significant estimates of the impact of duration to MM on service enrollment and 12 months; evidence consistent with impacts on STW and NSTW months in Phase 2, but conflicting evidence in Phase 1 and, especially, Phase 3; and little evidence of impacts on TWP completion. The evidence from Phase 2 combined with the conflicting nature of the Phase 3 evidence led us to obtain the selection date samples in order to better address the analysis limitations caused by the fact that SSA did not mail all Tickets as initially intended, especially during the later mail months for each phase. We also made some changes to the specification of the control variables, to improve the likelihood of detecting small impacts. These methodological improvements addressed the selection issue and improved estimator precisions, as intended. We also tried to assess how impacts varied with respect to the strength of the state labor market at the time the Ticket was mailed. This fine-tuning of the model specification did not change the overall nature of the findings.

APPENDIX C

DISCUSSION OF FINDINGS

This page has been left blank for double-sided copying.

We present further details on the results of the econometric analysis in this appendix. For the four event variables (service enrollment, beginning TWP, TWP completion, and STW), we show the linear model estimates of impacts of duration from rollout start to IMM along with the IV model estimates of impacts of duration from rollout start to MM. In addition, we also present the linear and IV model estimates of impacts on the number of NSTW months.

A. Clear Evidence of Impacts on Service Enrollment

1. Results from Linear Models with IMM

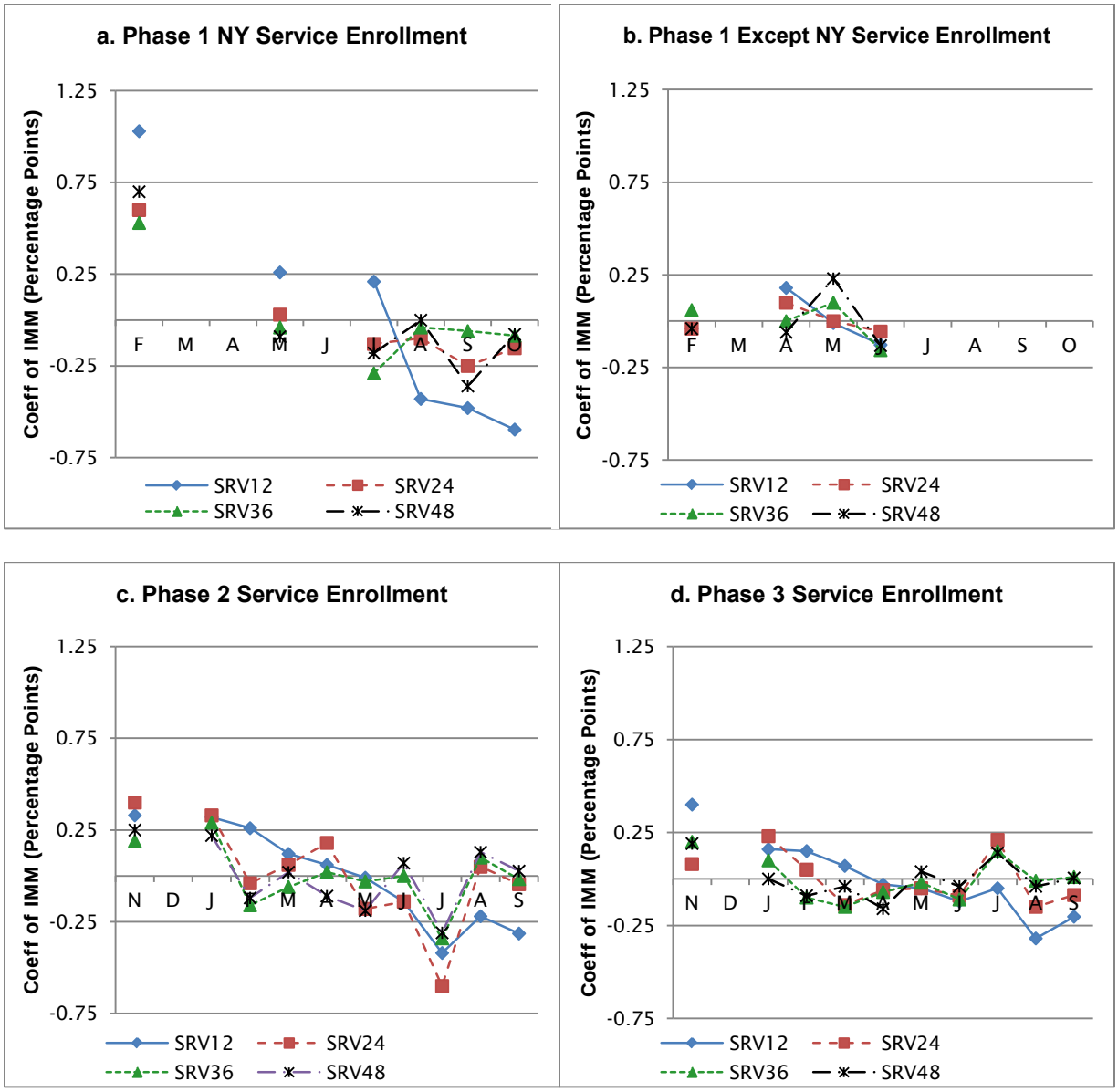
We found strongly significant evidence of negative effects of duration to intended mail month on service enrollment at 12 months following rollout start in Phase 1 NY, Phase 2, and Phase 3 (that is, the longer the duration, the lower the proportion enrolled), but no evidence of an effect in Phase 1 Except NY, where the duration of the rollout was relatively short. We found more limited evidence of impacts at 24, 36, and 48 months following rollout start. Consistent with expectations, this pattern implies that, on average, assignment of an early IMM accelerated the beneficiary's entry into service enrollment relative to assignment of a later IMM. It also appears that service enrollment for those with late IMM dates had, by the end of the observation period, largely caught up to enrollment for those with earlier dates.

In Exhibit C.1, we show the point estimates of the coefficients for IMM indicators from linear models of service enrollment at 12, 24, 36, and 48 months following the start of Ticket rollout in each phase. We plotted the coefficients over the full rollout period in each phase. In three of the four samples (excluding Phase 1 Except NY), the estimates at the first observation point (month 12) are clearly largest for the earliest intended mail months, and they decline as the value of the IMM approaches the last IMM, although not always monotonically. At later observation points (months 24, 36, and 48) there is much less evidence of a decline. For service enrollment at 12 months, the coefficients are jointly significant in the three samples (see joint significance test statistics in Appendix E, Table E1a). Thus, there is very significant, consistent evidence of impacts of duration to IMM on service enrollment at 12 months. For the Phase 2 sample, the coefficients are also jointly significant at 24, 36, and 48 months after the start of rollout, but in each case the decline with successive IMM is much smaller than at month 12, as indicated by the coefficient estimates shown in Exhibit C.1 for each mail month over the four observation points. For all the other phase samples, there is no statistically significant evidence of a continued effect after month 12. Although not significant, the point estimates at later observation points in Phase 1 except NY and in Phase 3 are consistent with a small, lingering effect at month 48, because those in the first IMM for each sample have a somewhat larger point estimate than for those in all later months.

2. Results from IV Models with MM

For the IV estimates of impacts of duration to the actual mail month, we focus on impacts as of month 12 following rollout start, and compare them to the estimates for the impact of duration to IMM, presented above. Also, because we did not find any evidence of impacts in the sample for Phase 1 Except NY, we discuss the IV estimates for the other three samples only (estimates for Phase 1 Except NY appear in Appendix E, Table E1a).

Exhibit C.1. Estimated Impacts of Duration to IMM on the Likelihood of Service Enrollment at 12, 24, 36, and 48 Months Following Rollout Start, by Phase



Notes: Coefficient estimates from regressions of service enrollment outcomes on IMM and control variables. Coefficients for each regression are constrained to sum to 0.0. SRV*mm* is the set of coefficients for impacts on service enrollment as of month *mm* after rollout start. The x-axis in each graph in this exhibit is labeled with the first letters of calendar months in each rollout phase. TTW was rolled out from February 2002 to October 2002 in Phase 1, from November 2002 to September 2003 in Phase 2, and from November 2003 to September 2004 in Phase 3.

In the first stage of IV estimation, the F-statistic for each endogenous MM is very large (not reported)—as anticipated given the strong correspondence we observed between the IMMs and the MMs (Exhibit II.3). As the first stage of IV estimation is the same for all outcome variables, this finding applies to the IV estimates for all of the outcome variables, not just service enrollment.

For each of the three samples considered here, we show in Exhibit C.2 plots the duration to MM coefficients from the IV analysis (right side) along with the corresponding earlier estimates

from the duration to IMM models 12 (left side). The band around each estimate is the estimate's 95 percent confidence interval. Results for joint tests of the hypothesis that all of the coefficients are zero appear at the bottom of each plot.

To a first approximation, the estimates appear to be arrayed tightly around a straight line. To facilitate discussion and to support the projections discussed in Appendix D, we fitted a line to the estimates by estimating a model in which we constrained the monthly coefficients to be on a straight line; that is, we assume that the marginal impact of duration to IMM (left side) or MM (right side) is constant over the rollout period. The slope of this line, its standard error, and the result for the test statistic for the tests of constant marginal impact also appear at the bottom of each figure.[39] The hypothesis was not rejected for the Phase 2 sample at the 12-month observation point (panels 2c and 2d), or at any observation points for the other samples, but was rejected for Phase 2 as of the later observation points at the 0.05 or 0.10 level, but not at the 0.01 level (Appendix E, Table E1a).

The IV estimates of the impacts of duration to MM are very similar to those from the linear models with IMM, and the confidence interval is also similar. The hypothesis that all of the IV coefficients are zero at 12 months is rejected at the 5 percent significance level by the chi-squared test for all three phase samples. Parallel to the linearity tests for duration to IMM, the statistical test of linearity for duration to MM was not rejected for service enrollment at 12 months in any sample, but was marginally rejected for Phase 2 estimates at 24, 36, and 48 months (see test statistics in Appendix E, Table E1a).

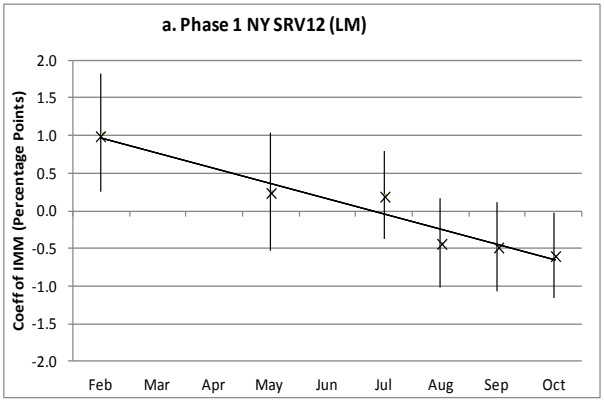
As anticipated, the IV estimates for duration to actual MM are somewhat larger in magnitude than the linear model estimates for duration to IMM, presumably because they account for the crossover effect of Tickets that were actually mailed in an earlier mail month than SSA originally intended. This can be seen most readily by comparing the estimated slope of the line for the restricted IV coefficients (right side) with the corresponding slope for the linear model estimate (left side); in each case, the former is more negative than the latter. The differences in magnitude are very small, however, likely reflecting the fact that a large majority of Tickets was mailed in the intended mail month.

As documented in the appendix, the IV estimates for service enrollment as of later observation points also suggest that the relationship between duration to MM and service enrollment is slightly stronger than the relationship between duration to IMM and service enrollment. The differences are so slight that our conclusion is analogous to that for the impacts of duration to IMM on service enrollment: by the end of the observation period, service enrollment for those with MM dates late in the rollout had largely caught up to service enrollment for those with earlier MM dates.

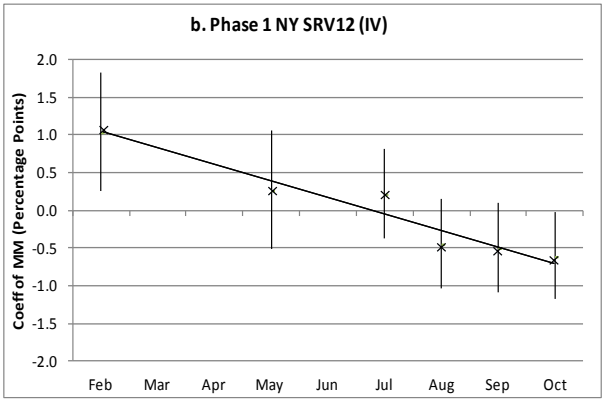
Extrapolation of the trend lines for the IV estimates in each phase to 12 months provides an estimate of the impact on service enrollment of mailing the Ticket in the first rollout month versus not mailing the Ticket until month 13 or later: 2.5 percentage points for Phase 1 NY, 1.0 percentage

[39] The degrees of freedom in the denominator of the F-test is the number of observations for the NY sample and the number of clusters (50 states, District of Columbia, Puerto Rico, and other U.S. Territories) for the Phase 2 and Phase 3 samples. See Section B in Chapter III for more technical discussion.

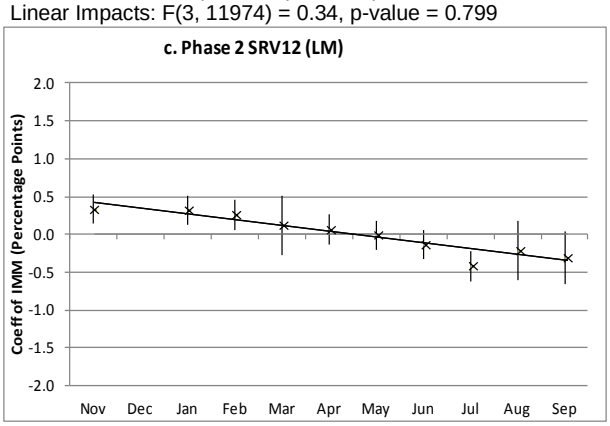
Exhibit C.2. Estimated Impacts of Duration to IMM (left) and MM (right) on the Likelihood of Service Enrollment at 12 Months Following Rollout Start in Phase 1 NY, Phase 2 and Phase 3



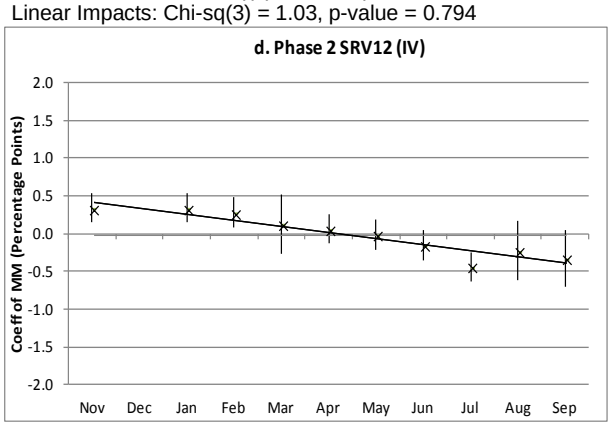
Slope of Line = -0.0021, SE = 0.000648, p-value = 0.001
Zero Coefficients: F(5, 11974) = 2.33, p-value = 0.040
Linear Impacts: F(3, 11974) = 0.34, p-value = 0.799



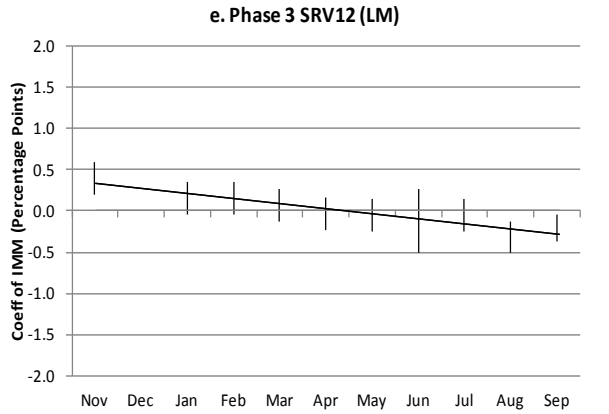
Slope of Line = -0.00215, SE = 0.00066, p-value = 0.001
Zero Coefficients: Chi-sq(5) = 11.72, p-value = 0.039
Linear Impacts: Chi-sq(3) = 1.03, p-value = 0.794



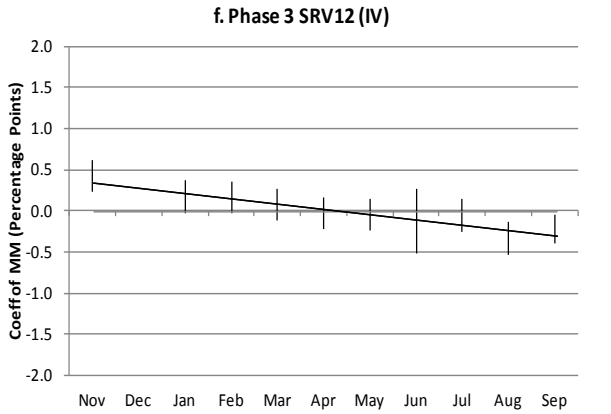
Slope of Line =-0.00078, SE = 0.000164, p-value <0.001
Zero Coefficients: F(9, 53) = 5.19, p-value < 0.001
Linear Impacts: F(7, 53) = 0.87, p-value = 0.533



Slope of Line =-0.00084, SE =0.00017, p-value < 0.001
Zero Coefficients: Chi-sq(9) = 49.68, p-value < 0.001
Linear Impacts: Chi-sq(7) = 6.50, p-value = 0.482



Slope of Line =-0.00060, SE = 0.000124, p-value < 0.001
Zero Coefficients: F(9, 53) = 5.36, p-value < 0.001
Linear Impacts: F(7, 53) = 0.87, p-value = 0.538



Slope of Line = -0.00066 SE = 0.00013, p-value < 0.001
Zero Coefficients: Chi-sq(9) = 51.2, p-value = 0.001
Linear Impacts: Chi-sq(7) = 6.02, p-value = 0.538

Notes: Estimates in left panels are regression coefficients from the model with IMM. Estimates on the right are IV estimates from the model with MM, using IMM as instruments. All coefficients are constrained to sum to zero. The slope of the linear trend line reflects the estimates from the same model with linear restrictions imposed on the coefficients; the slope is reported at the bottom of the panel, in percentage points. The first test statistic at the bottom of each panel is for the test of the null hypothesis that all the coefficients are zero, and the second is for the linear restrictions on the coefficients.

points for Phase 2, and 0.8 percentage points for Phase 3.[40] Although the relatively large value for Phase 1 NY suggests that the impact in NY was larger than in Phase 2 or 3, it might also be that the difference is due to sampling error—much higher for the NY sample than for the other two because of the relatively small sample size. Given the relatively narrow confidence intervals for Phase 2 and 3, it seems clear that the impacts for Phase 2 and 3 were quite similar. In Appendix D, we use this result along with results for the later observations points to project impacts of mailing a Ticket in the first rollout month versus not mailing a Ticket as of 24, 36 and 48 months after rollout start.

B. Unclear Evidence of Impacts on TWP Start and Completion

As indicated earlier, we also analyzed the impact of duration to mail month on TWP start and TWP completion at 12, 24, 36, and 48 months following rollout start in each phase. Impacts on TWP start or completion could occur even if impacts on STW do not occur, because Ticket might induce beneficiaries to work and earn enough to complete the TWP without earning enough to have their benefits suspended for work. One possibility is that the knowledge Ticket users gain about how earnings affect their benefits helps them work or earn more without losing their benefits.

Almost all estimates of impacts on TWP start—for every sample and every observation point—were statistically insignificant, and we do not consider them further (all estimates are reported in Appendix E, Table E1b). Estimates of impacts on TWP completion are more often significant, but present a mixed picture.[41]

1. Results from Linear Models with IMM

In Exhibit C.3, we present the estimated coefficients from linear models for impacts of duration to IMM on TWP completion as of 12, 24, 36, and 48 months following rollout. The coefficients in each phase do not clearly decline with duration to IMM. For all four follow-up periods, the coefficients were not jointly statistically different from zero for either of the two Phase 1 samples (statistics shown in Appendix E, Table E1c). For Phases 2 and 3, the coefficients were jointly significant at 24, 36, and 48 months (at 5 percent level), but not at 12 months—consistent with the fact that it takes at least 9 consecutive months to complete the TWP. We discuss the estimates for the three of the four samples further after presenting the IV estimates in the next section; we do not consider the Phase 1 Except NY sample further due to the lack of any pattern for the linear model estimates and the lack of evidence of impacts on service enrollment.

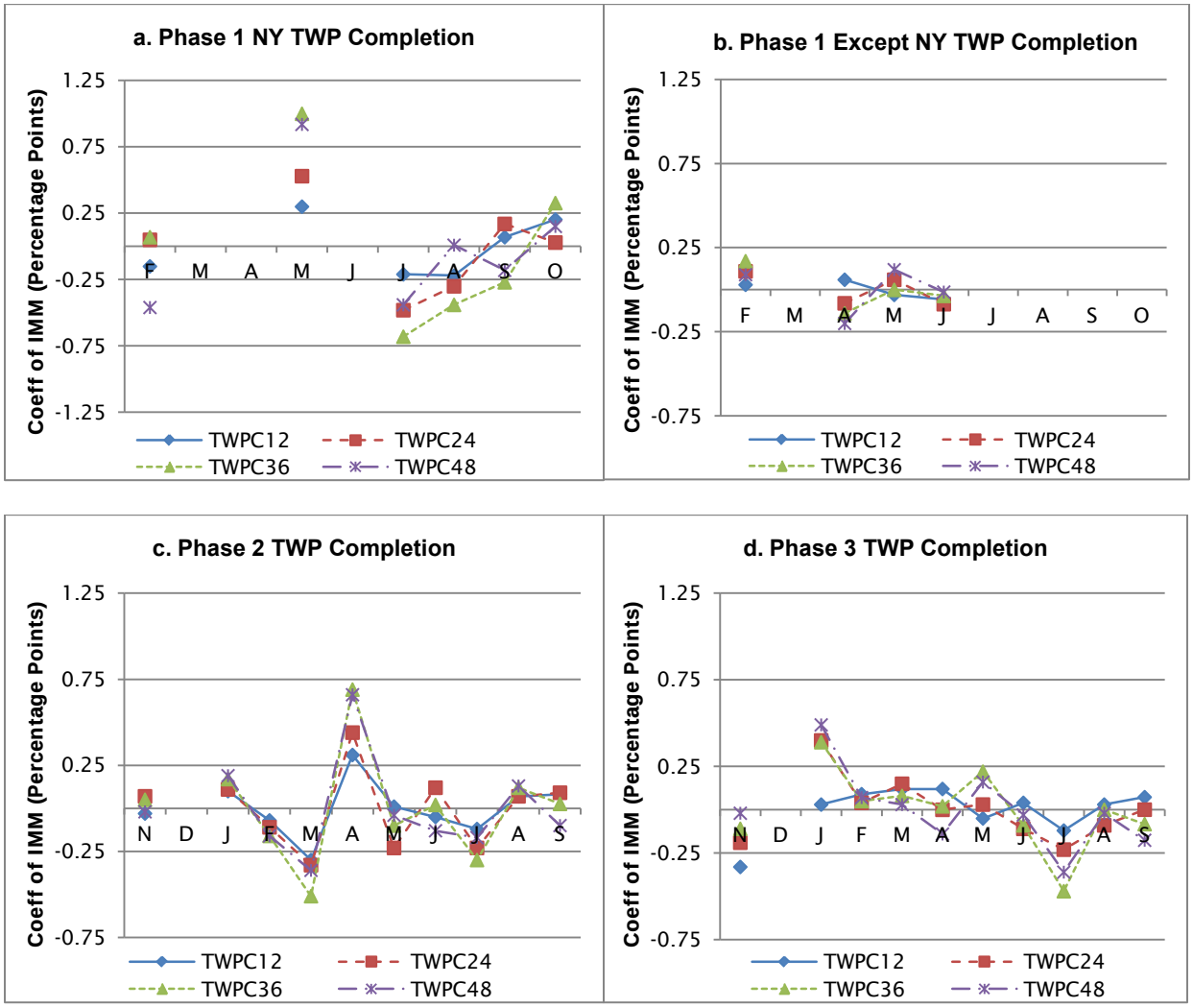
2. Results from IV Models with MM

For TWP completion, Exhibit C.4 plots the estimated coefficients from the linear and IV models for impacts on the likelihood of TWP completion at 48 months after the start of rollout, along with their 95 percent confidence intervals and estimated lines with linear restrictions imposed; results at 24 and 36 months for both phases are quite similar. Statistics from the joint tests of the

[40] These values were obtained by multiplying the slopes of the respective fitted lines by 12.

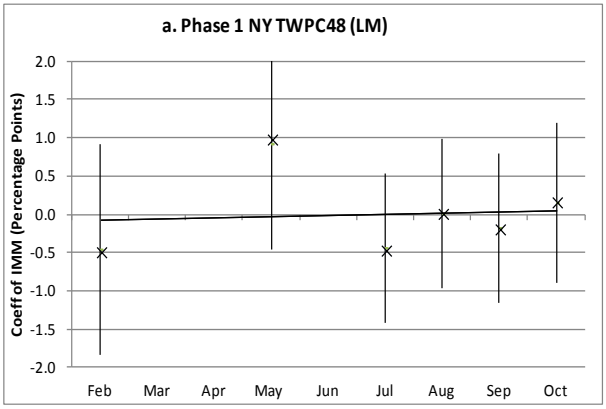
[41] One factor that might influence the estimates for TWP start is that TWP start is underreported in the administrative data. Typically we have found that about 20 percent of those with a TWP completion date do not have a TWP start date. This likely reflects the fact that the TWP start date has no immediate consequence for benefits due, whereas TWP completion date is a critical marker for purposes of determining future benefit payments.

Exhibit C.3. Estimated Impacts of Duration to IMM on TWP Completion at 12, 24, 36, and 48 Months Following Rollout Start, by Phase

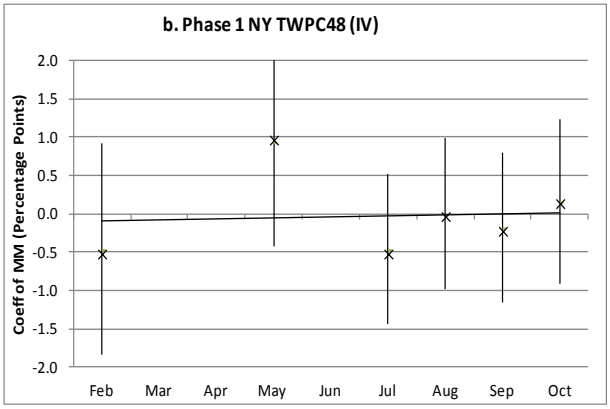


Notes: Coefficient estimates from regressions of service enrollment outcomes on IMM and control variables. Coefficients for each regression are constrained to sum to 0.0. TWPC*mm* is the set of coefficients for impacts on TWP completion as of month *mm* after rollout start. The x-axis in each graph in this exhibit is labeled with the first letters of the calendar months in each rollout phase. TTW was rolled out from February 2002 to October 2002 in Phase 1, from November 2002 to September 2003 in Phase 2, and from November 2003 to September 2004 in Phase 3.

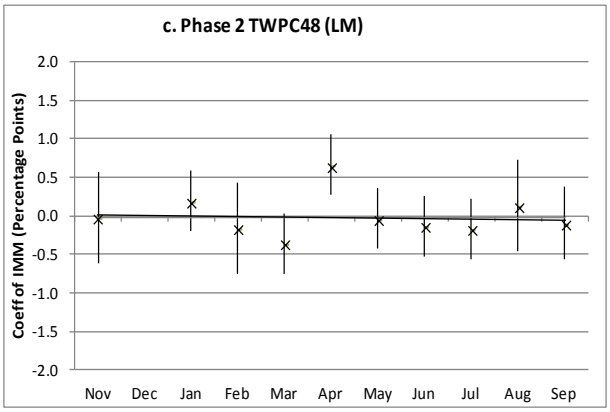
Exhibit C.4. Estimated Impacts of Duration to IMM (left) and MM (right) on TWP Completion at 48 Months Following Rollout Start in Phase 1 NY, Phase 2, and Phase 3



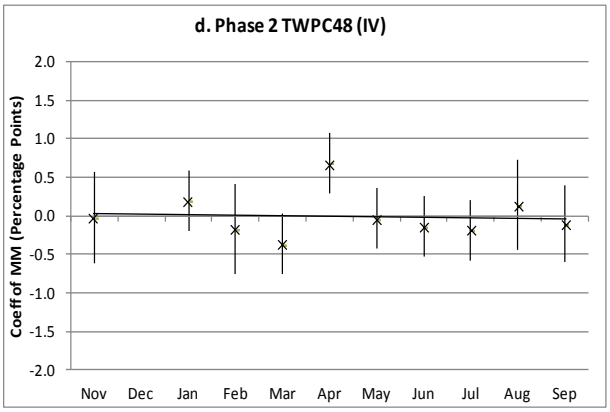
Slope of Line =0.00016, SE = 0.00107, p-value =0.878
Zero Coefficients: F(5, 11974) = 0.47, p-value = 0.798
Linear Impacts: F(3,11974) = 0.72, p-value = 0.537



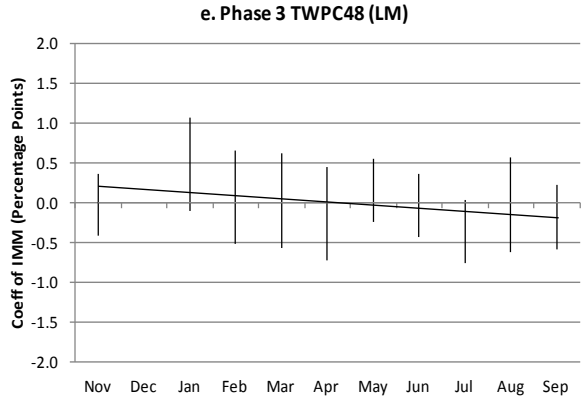
Slope of Line =0.00017, SE = 0.00109, p-value = 0.875
Zero Coefficients: Chi-square(5) = 2.37, p-value = 0.797
Linear Impacts: Chi-square(3) = 2.18, p-value = 0.536



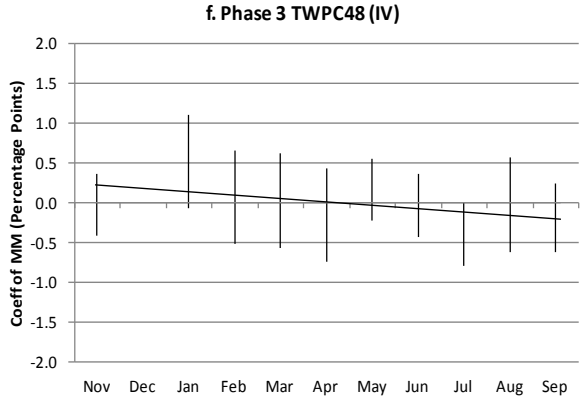
Slope of Line =-0.00006, SE = 0.00032, p-value =0.843
Zero Coefficients: F(9, 53) = 2.36, p-value=0.025
Linear Impacts: F(7, 53) = 2.83, p-value=0.014



Slope of Line =-0.00007, SE = 0.00034, p-value = 0.842
Zero Coefficients: Chi-square(9) = 21.11, p-value = 0.012
Linear Impacts: Chi-square(7) = 19.75, p-value = 0.006



Slope of Line =-0.00038, SE = 0.00018, p-value = 0.038
Zero Coefficients: F(9, 53) = 2.43, p-value = 0.021
Linear Impacts: F(7, 53) = 1.83, p-value = 0.102



Slope of Line =-0.00041, SE = 0.00019, p-value = 0.03
Zero Coefficients: Chi-square(9) = 23.02, p-value = 0.006
Linear Impacts: Chi-square(7) = 12.95, p-value = 0.073

Notes: Estimates in left panels are regression coefficients from the model with IMM. Estimates on the right are IV estimates from the model with MM, using IMM as instruments. All coefficients are constrained to sum to zero. Lines are for the same models with linear restrictions imposed on the coefficients; the slope is reported at the bottom of the panel, in percentage points. The first test statistic at the bottom of each panel is for the test of the null hypothesis that all the coefficients are zero, and the second is for the linear restrictions on the coefficients.

hypothesis that all of the coefficients are zero are reported at the bottom of each plot. Again, the linear model and IV model estimates and confidence intervals for each sample are very similar to each other.

For the Phase 1 NY sample (panels 4a and 4b), there is clearly no evidence of an effect of duration to either IMM or MM on TWP completion as of month 48. For both Phase 2 (panels 4c and 4d) and 3 (panels 4e and 4f), the estimates of an impact of duration to either IMM or MM on TWP completion are jointly significant at the 5 percent level. Despite the significance of the Phase 2 estimates, however, the evidence is not indicative of an impact of duration to MM because of the pattern of the coefficients. The hypothesis of zero coefficients is rejected because the March coefficient is very low and the April coefficient is very high, but their values relative to each other are inconsistent with the hypothesis that duration to MM reduces TWP completion. The slope of the line for the constrained estimates is very small and not statistically significant, and we reject the hypothesis that the marginal effect of a one-month delay in Ticket mailing is independent of the rollout month (Appendix E, Table E1c).

The Phase 3 pattern of coefficients (panel 4f) is more consistent with the hypothesized negative impact on TWP completion. The slope of the line for the restricted estimates is steeper than for Phase 2 (panel 4d). Extrapolation of the line to 12 months indicates that a 12-month delay in mailing the Tickets reduced the likelihood of TWP completion over the 48-month period by 0.5 percentage points—a substantial reduction relative to the 12-month estimated impact on service enrollment (-0.8 percentage points). It is noteworthy, however, that, as in Phase 2, the null hypothesis of no impact is rejected again because of just two coefficients with relatively large magnitudes—one positive and one negative; the fitted line is steeper than in Phase 2 because the difference in coefficients for these two months is in the expected direction and they are further apart from each other. Overall, however, the pattern of coefficients in Phase 3 is just as irregular as in Phase 2, and the results just as clearly reject the hypothesis that the marginal effect of a one-month delay in Ticket mailing is independent of the rollout month (Appendix E, Table E1c).

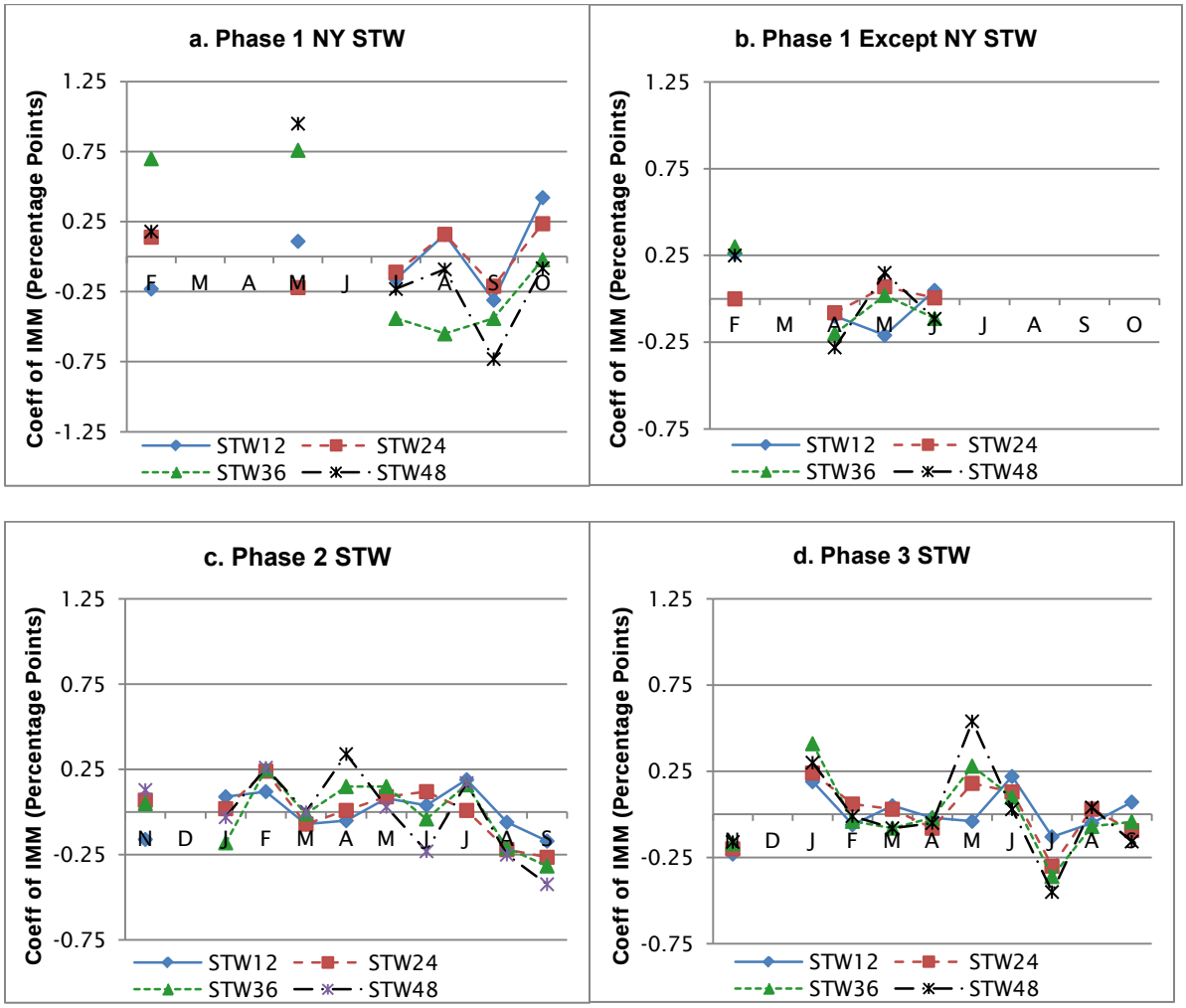
In summary, the linear and IV model estimates are consistent with negative impacts of duration to mail month on TWP completion in Phase 3, but not in the other phases. The findings for Phases 2 and 3 are each driven by coefficients for two months that are large relative to the coefficients for all other months—one positive and the other negative—and the location of these months within the respective rollout periods differs in a manner that suggests no impact in Phase 2 and a substantive impact for Phase 3. In both samples, the hypothesis of a constant marginal impact of duration to MM is clearly rejected. The patterns of the statistics leave open the possibility that the estimates for both phases are due to chance rather than evidence that duration to MM has a negative impact on TWP completion.

C. Unclear Evidence of Impacts on STW

1. Results from Linear Models with IMM

For STW, we found weak evidence that early Ticket mailing accelerated attainment of STW for Phase 2, but no such evidence for the other phases. In Exhibit C.5, we present the estimated coefficients from linear models for the impact of duration to IMM on STW as of 12, 24, 36, and 48 months following rollout start. Unlike the corresponding impact estimates for service enrollment (Exhibit C.1), none of these estimates displays patterns that are clearly consistent with an impact of duration to IMM on STW as of the outcome month.

Exhibit C.5. Estimated Impacts of Duration to IMM on the Likelihood of STW at 12, 24, 36, and 48 Months Following Rollout Start, by Phase



Notes: Coefficient estimates from regressions of service enrollment outcomes on IMM and control variables. Coefficients for each regression are constrained to sum to 0.0. STW*mm* is the set of coefficients for impacts on STW as of month *mm* after rollout start. The x-axis in each graph in this exhibit is labeled with the first letters of the calendar months in each rollout phase. TTW was rolled out from February 2002 to October 2002 in Phase 1, from November 2002 to September 2003 in Phase 2, and from November 2003 to September 2004 in Phase 3.

The joint test statistics (shown in Appendix E, Table E1d) indicate that estimated coefficients for the Phase 1 NY sample (panel 5a) were not jointly significantly different from zero at the 5 percent level for any of the follow-up periods. Although that is surprising, given the large and significant estimates for duration to service enrollment for this sample, keep in mind that this sample is relatively small, so confidence intervals are relatively wide. Also somewhat surprising, for Phase 1 Except NY (panel 5b), the coefficients for STW at 12 months are jointly significant at the 5 percent level only, although not for coefficients for the later months. Again, this might reflect the relatively small sample.

For the Phase 2 sample (panel 5c), we also find that the coefficients from the linear models were not jointly statistically different from zero at the 5 percent level as of any of the follow-up months. At 48 months, they are significant at the 10 percent level, however. For the Phase 3 sample (panel 5d), the coefficients are jointly significant at the 5 percent level for STW as of each of the

follow-up months. However, the pattern of the coefficients, which we will consider further in the presentation of the IV estimates, is inconsistent with the hypothesis that duration to IMM had a negative impact on attaining STW over the rollout period.

2. Results from IV Models with MM

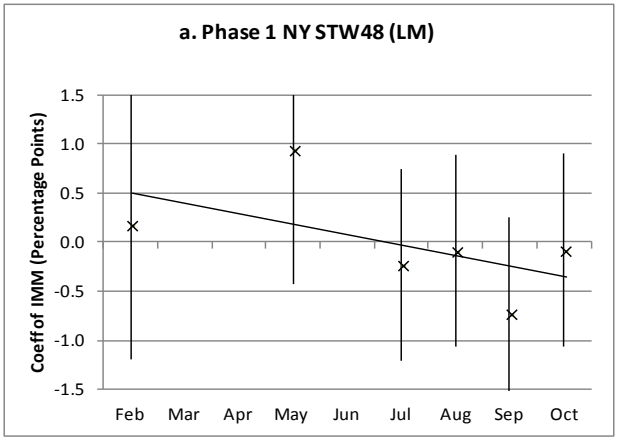
Because it might take a beneficiary a long time to achieve STW after service enrollment (for service delivery, attaining a job with sufficiently high earnings, and completing the TWP and grace period), STW impacts are more likely to be observed as of the later observation months than as of month 12. We focus on the STW results at month 48; although they provide only weak evidence of an impact, the evidence of an impact as of this month is stronger than in the three earlier observation months. As earlier, we ignore the Phase 1 Except NY results due to the short duration of the rollout period and lack of an evidence of impact even on service enrollment (estimates for Phase 1 Except NY are shown in Appendix E, Table E1d).

In Exhibit C.6, we show plots for the second-stage estimated coefficients from the IV analysis (right side) along with the corresponding estimates from the linear models for STW as of month 48 (left side). As in Exhibit C.2, each plot also shows confidence intervals for the point estimates and the line estimated by imposing linear restrictions on the coefficients.

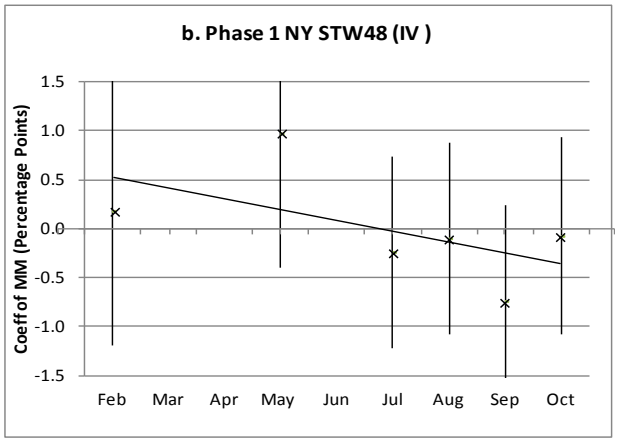
Once again, the linear model (left side) and IV model (right side) estimates and confidence intervals for each sample are very similar to each other. The linear trend line for each plot indicates a clear negative slope for the Phase 1 NY (panels 6a and 6b) and Phase 2 samples (panels 6c and 6d), but a very small negative slope for Phase 3 (panels 6e and 6f). The Phase 1 NY sample estimates are not jointly significant, reflecting the wide confidence intervals for each estimate. As with the linear models, the hypothesis that all of the IV coefficients are zero is rejected at the 5 percent significance level by the chi-squared test only for the Phase 3 sample, and at the 10 percent level for Phase 2; for the Phase 3 estimates, however, it is important to note that the slope of the line under the linear restrictions is very small—much smaller than for Phase 2. The reason that the joint hypothesis of zero coefficients is rejected, despite this negative slope, is that the coefficients for two months (May and July) are both significantly different from zero. Unlike with the service enrollment estimates at 12 months, we do not see a fairly smooth declining pattern of coefficients from the first mail month to the last. This irregular pattern is reflected in the very large value of the test statistics for the hypothesis the marginal effect of delaying the mailing by one month is independent of the MM (see Appendix E, Table E1d).

Although the hypothesis that all coefficients in the IV estimates for Phase 2 are zero is only rejected at the 10 percent significance level, the pattern of the coefficients for Phase 2 is much more consistent with an impact of duration to MM on STW than are the corresponding patterns for Phase 1 NY and Phase 3. We also note that the slope of the line for the restricted Phase 2 IV estimates (panel 6d) is slightly steeper than the corresponding slope for duration to IMM (panel 6c), as we would expect if the IV method corrects for the effect of Tickets that were mailed in months other than the IMM. Extrapolating the fitted line to 12 months, the point estimates suggest that a 12-month delay in mailing the Ticket decreased attainment of STW at 48 months by about

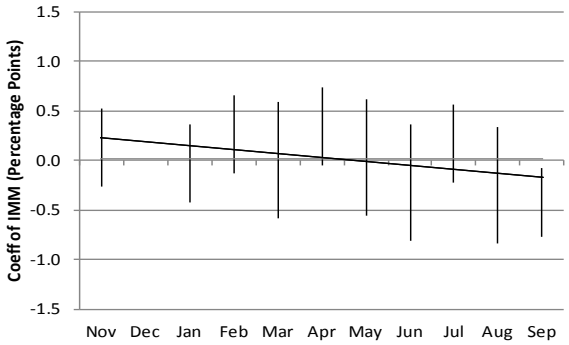
Exhibit C.6. Estimated Impacts of Duration to IMM (left) and MM (right) on STW at 48 Months Following Rollout Start in Phase 1 NY, Phase 2 and Phase 3



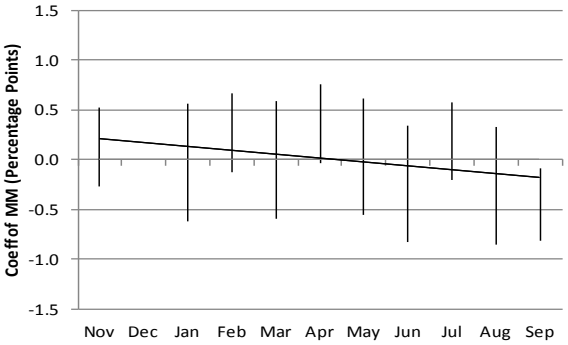
Slope of Line =-0.00107, SE = 0.00101, p-value = 0.29
Zero Coefficients: F(5, 11974) = 0.70, p-value = 0.622
Linear Impacts: F(3, 11974) = 0.62, p-value = 0.600



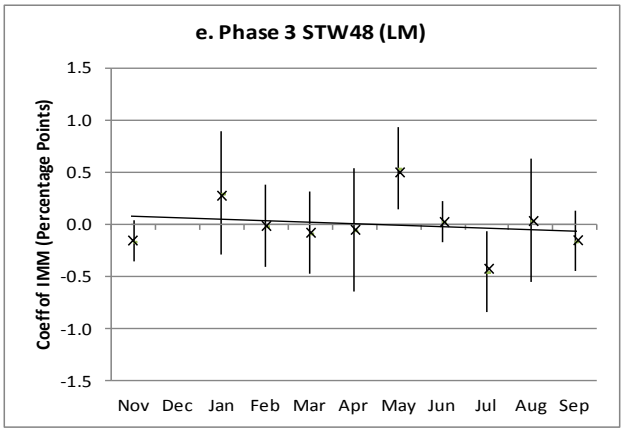
Slope of Line =-0.0011, SE = 0.00103, p-value = 0.29
Zero Coefficients: Chi-sq(5) = 3.52, p-value = 0.620
Linear Impacts: Chi-sq(3) = 1.89, p-value = 0.595



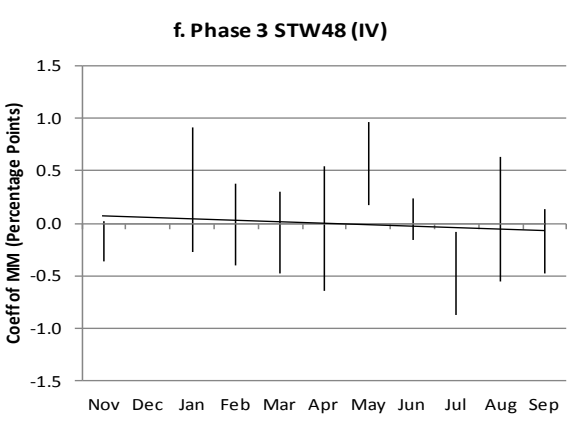
Slope of Line =-0.00044, SE = 0.00024, p-value = 0.075
Zero Coefficients: F(9, 53) = 1.78, p-value = 0.095
Linear Impacts: F(7, 53) = 0.73 p-value = 0.648



Slope of Line =-0.00047, SE = 0.00026, p-value = 0.068
Zero Coefficients: Chi-sq(9) = 16.29, p-value = 0.061
Linear Impacts: Chi-sq(7) = 5.19 p-value = 0.636



Slope of Line =-0.00015, SE = 0.00018, p-value = 0.429
Zero Coefficients: F(9, 53) = 3.606, p-value = 0.001
Linear Impacts: F(7, 53) = 3.06, p-value = 0.009



Slope of Line =-0.00016, SE = 0.00020, p-value = 0.433
Zero Coefficients: Chi-sq(9) = 33.00, p-value < 0.001
Linear Impacts: Chi-sq(7) = 22.28, p-value = 0.002

Notes: Estimates in left panels are regression coefficients from the model with IMM. Estimates on the right are IV estimates from the model with MM, using IMM as instruments. All coefficients are constrained to sum to zero. Lines are for the same models with linear restrictions imposed on the coefficients; the slope is reported at the bottom of the panel, in percentage points. The first test statistic at the bottom of each panel is for the test of the null hypothesis that all the coefficients are zero, and the second is for the linear restrictions on the coefficients.

0.6 percentage points.[42] The corresponding estimate for STW at month 48 in Phase 3 is less than 0.2 percentage points.

Although the results for Phase 2 are consistent with a substantial negative effect, they are not statistically strong. The Phase 1 NY and Phase 3 results provide very weak support, at best. The Phase 2 and 3 results for STW also seem inconsistent with the Phase 2 and 3 results for TWP completion. As reported earlier, the point estimate for the impact on TWP completion is roughly five times larger in Phase 3 than in Phase 2 (based on the slope of the line when the linear restrictions are imposed), but the point estimate for the impact on STW is about three times larger in Phase 2 than in Phase 3. In other words, the estimates suggest that for Phase 3, a substantial impact on TWP completion did not translate into a substantial impact on STW, whereas in Phase 2 a much more modest impact on TWP completion translated into a substantial impact on attainment of STW. A possible explanation of these inconsistent results is that they are all due to chance.

D. Unclear Evidence of Impacts on NSTW Months

1. Results from Linear Models with IMM

In Exhibit C.7, we present the estimated coefficients from linear models for impacts of duration to IMM on NSTW months as of 12, 24, 36, and 48 months following rollout. The coefficients in each phase do not clearly decline with duration to IMM. For all four follow-up periods, the coefficients were not jointly statistically different from zero at the 5 percent level for either of the two Phase 1 samples (statistics shown in Appendix E, Table E1e). For Phase 2, the coefficients were jointly significant at 36 months (at the 5 percent level) and at 48 months (at the 10 percent level). For Phase 3, the coefficients were jointly significant for NSTW months at 12, 24, and 48 months. We further discuss the estimates after presenting the IV estimates in the next section, but do not consider the Phase 1 Except NY sample further due to the lack of any pattern for the linear model estimates and the lack of evidence of impacts on service enrollment for this phase.

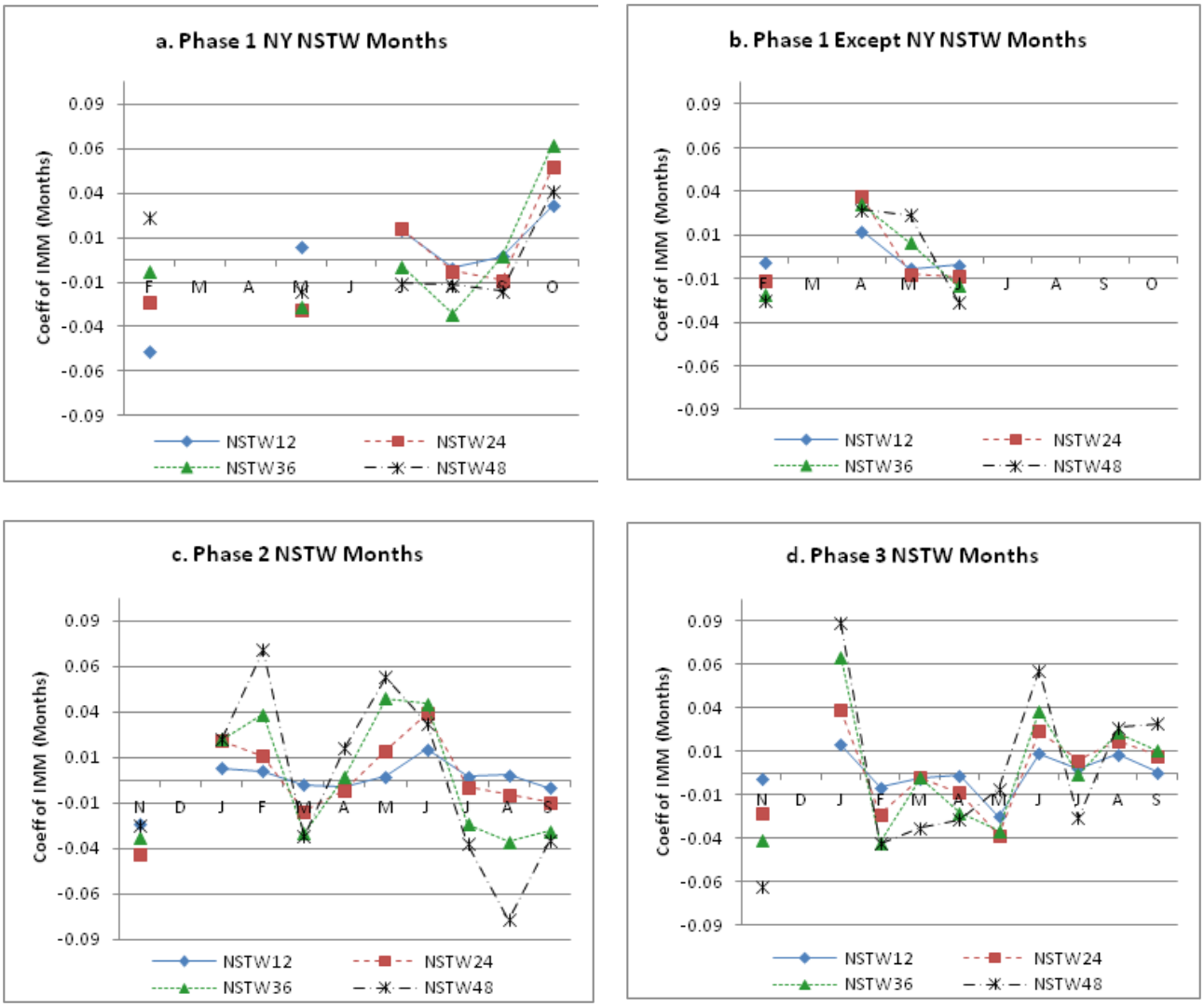
2. Results from IV Models with MM

In Exhibit C.8, we plot the estimated coefficients from the linear (left side) and IV (right side) models for impacts on the number of NSTW months at 48 month after rollout started, along with their 95 percent confidence intervals and a line representing the estimates when linear restrictions are imposed.

For the Phase 1 NY sample (panels 8a and 8b), there is clearly no evidence of an effect of duration to either IMM or MM on NSTW months as of month 48. For Phase 2, the IV estimates for duration to MM are jointly significant at the 5 percent level (panel 8d); the linear model estimates are significant only at the 10 percent level (panel 8c). Extrapolating the fitted line to 12 months suggests that a 12-month delay in distribution of the Tickets reduces the number of NSTW months over the 48-month period by an average of 0.07 months. Similar to the Phase 2 results for STW, the Phase 2 results for NSTW months are consistent with a modest negative impact of duration to MM on service enrollment. We do not reject the null hypothesis that the impact of duration to MM is linear over the rollout period.

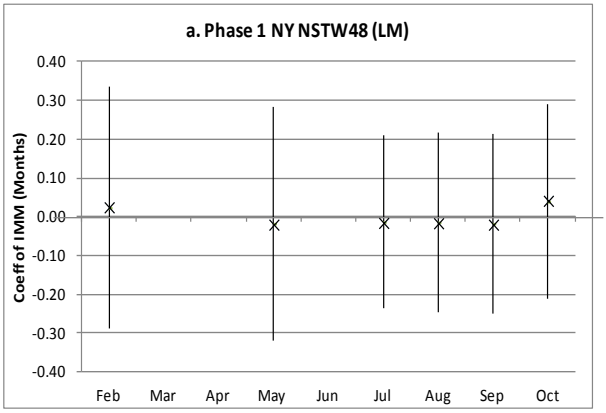
[42] This figure was obtained by multiplying the coefficient of the fitted line by 12.

Exhibit C.7. Estimated Impacts of Duration to IMM on NSTW Months at 12, 24, 36, and 48 Months Following Rollout Start, by Phase

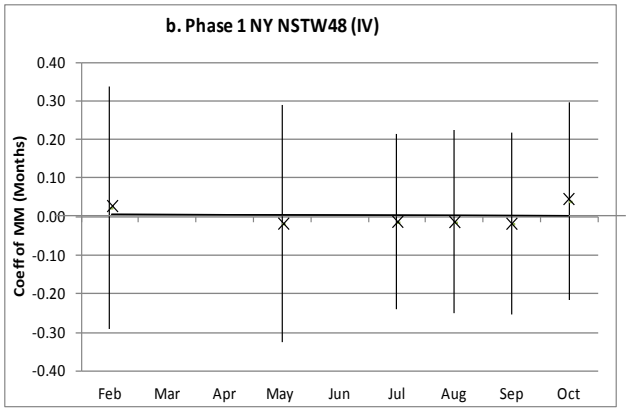


Notes: Coefficient estimates from regressions of service enrollment outcomes on IMM and control variables. Coefficients for each regression are constrained to sum to 0.0. NSTW*mm* is the set of coefficients for impacts on NSTW months as of month *mm* after rollout start. The x-axis in each graph in this exhibit is labeled with the first letters of the calendar months in each rollout phase. TTW was rolled out from February 2002 to October 2002 in Phase 1, from November 2002 to September 2003 in Phase 2, and from November 2003 to September 2004 in Phase 3.

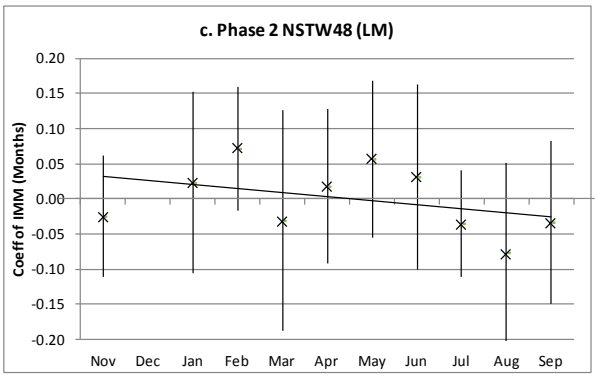
Exhibit C.8. Estimated Impacts of Duration to IMM (left) and MM (right) on NSTW Months at 48 Months Following Rollout Start in Phase 1 NY, Phase 2 and Phase 3



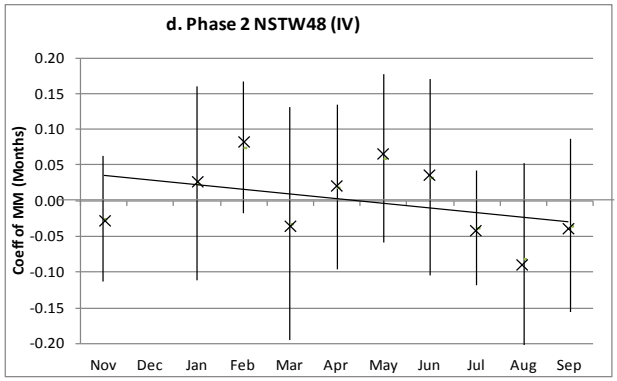
Slope of Line =0.00113, SE = 0.02485, p-value =0.964
Zero Coefficients: F(5, 11974) = 0.03, p-value > 0.999
Linear Impacts: F(3, 11974) = 0.00, p-value > 0.999



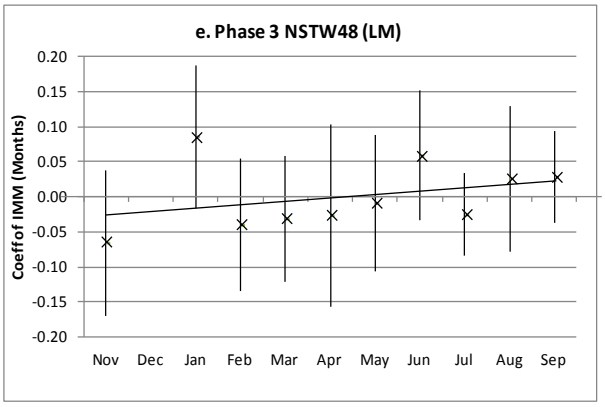
Slope of Line =0.00111, SE =0.02542, p-value = 0.965
Zero Coefficients: Chi-square(5) = 0.15, p-value > 0.999
Linear Impacts: Chi-square(3) = 0.01, p-value > 0.999



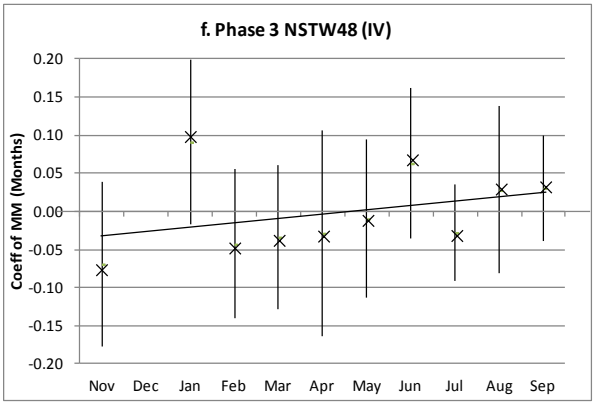
Slope of Line =-0.00565, SE = 0.00572, p-value =0.328
Zero Coefficients: F(9, 53) = 1.98, p-value = 0.060
Linear Impacts: F(7, 53) = 1.26, p-value = 0.290



Slope of Line =-0.00595, SE = 0.00603, p-value = 0.324
Zero Coefficients: Chi-square(9) = 17.85, p-value = 0.037
Linear Impacts: Chi-square(7) = 9.04, p-value = 0.250



Slope of Line = 0.00503, SE = 0.00502p-value =0.321
Zero Coefficients: F(9, 53) = 2.37, p-value = 0.025
Linear Impacts: F(7, 53) = 1.95, p-value = 0.080



Slope of Line = 0.00549, SE = 0.00541, p-value = 0.309
Zero Coefficients: Chi-square(9) = 22.88, p-value = 0.006
Linear Impacts: Chi-square(7) = 14.43, p-value = 0.044

Notes: Vertical scale for top panels are twice is half as large as for middle and lower panels. Estimates in left panels are regression coefficients from the model with IMM. Estimates on the right are IV estimates from the model with MM, using IMM as instruments. All coefficients are constrained to sum to zero. Lines are for the same models with linear restrictions imposed on the coefficients; the slope is reported at the bottom of the panel, in months. The first test statistic at the bottom of each panel is for the test of the null hypothesis that all the coefficients are zero, and the second is for the linear restrictions on the coefficients.

For Phase 3, the linear (panel 8e) and IV (panel 8f) model estimates are also quite similar, and we reject the hypothesis that all of the Phase 3 IV coefficients are zero at the 5 percent significance. However, the pattern of the coefficients is indicative of *positive* effects, rather than negative effects—as illustrated by the positive slopes of the fitted lines.

In summary, as with the STW results, the NSTW-months results for Phase 2 are consistent with the hypothesis of small negative impacts of duration to IMM or MM. The results are not statistically strong, however, and the conclusion that the impact was negative is undermined by the absence of any evidence consistent with negative impacts in the Phase 1 NY and Phase 3 samples as well as the inconsistent evidence of impacts on TWP completion in the Phase 2 and 3 samples. The analysis of total impacts, presented in Appendix D, reinforces the conclusion that the marginally significant impacts on STW and NSTW months found for Phase 2 and simply the result of chance.

This page has been left blank for double-sided copying.

APPENDIX D

DISCUSSION OF PROJECTIONS FOR TOTAL IMPACTS

This page has been left blank for double-sided copying.

Estimates for the impact of the duration to MM on any outcome at later points in time can be converted to projections of "total impacts" as of those later points if certain assumptions are maintained. That is, the projections are estimates of cumulative impacts as of that later month of mailing the Ticket in the first rollout month versus not mailing it at all. We call these estimates projections because of their reliance on the maintained assumptions.

Given the statistical significance and magnitude of estimated impacts for duration to MM on service enrollment, it seems reasonable to construct estimates of total impacts for that outcome. We first describe the conceptual approach, using Phase 2 estimates for illustrative purposes. We then consider the service enrollment estimates for all four samples. They show that the total impact of mailing the Ticket in the first rollout month on service enrollment at month 12 increases to successive larger values at months 24, 36, and 48. Findings for all other outcomes—TWP start, TWP completion, STW and NSTW months—are clearly different: there is no evidence of a substantial total impact as of 48 months; nor does such evidence emerge at 12, 24 and 36 months.

A. Conceptual Approach

To understand the conceptual approach, it is helpful to first consider a more simple approach that is clearly flawed: extrapolation to 24, 36, and 48 months of the fitted line for the restricted IV estimate of the impact of duration to MM on service enrollment at 12 months. As seen in Appendix C (Exhibit C.2), the unrestricted IV estimates for the monthly coefficients for service enrollment at 12 months follow an approximately linear pattern in all samples; they show no indication that the effect of a one-month delay in mailing a Ticket changes as the last mail month is passed. If we assume the effect continues to be linear through any later month, then we could simply use the fitted line for the impact of duration to MM on service enrollment as of month 12 to project enrollment in the later month, including month 48. Eventually, though, the marginal impact must decline; otherwise the total impact would be unlimited. Further, while it is plausible that the impact is linear for the first year or so following the rollout start, it is also plausible that it would start to decline well before the end of our observation period—month 48. As will be seen, a less restrictive approach that makes more use of the data than the simple extrapolation leads to a projection that is consistent with such a decline.

The projection methodology we used relies on two assumptions. The first is that the marginal impact of delaying the mailing of the Ticket on service enrollment (or other outcome) as of month 12, 24, 36 or 48 is linear through month 13 of the 48-month observation period for each sample (hereafter the "linearity" assumption). That is, we extrapolate the fitted line to month 13 only, but do so at each observation point. This assumption can be tested over the duration of the rollout period. For Phases 2 and 3, the rollout period is only three months shorter than 13 months; for each phase the last rollout month is month 10 of the observation period. In fact, we previously presented the results of such tests (Appendix C, Section A). For Phase 2, we failed to reject the linearity hypothesis for service enrollment as of month 12, but marginally rejected it for enrollment as of months 24, 36 and 48. For Phase 1 NY and Phase 3, we failed to reject linearity for service enrollment as of each observation month.

Although the linearity assumption requires some extrapolation beyond the rollout period, it requires much less extrapolation than use of extrapolation alone to produce total impact estimates in month 24 and beyond. This approach is somewhat restrictive in that it assumes linearity over a 13-month period, but it allows the data to determine how delays in mailing Tickets within that period affect outcomes at each of four points: 12, 24, 36 and 48 months.

Given the linearity assumption, for Phase 2 the point estimate of the impact on service enrollment as of month 12 of mailing the Ticket in month 1 instead of month 13 is the negative of 12 times the slope of the fitted IV line for service enrollment at month 12—1.00 (= -12 x 0.083). Put differently, mailing the Ticket in month 1 instead of month 13 increases cumulative enrollment at month 12 by 1.00 percentage points. Applying the same approach to service enrollment as of 24, 36 and 48 months implies that mailing the Ticket in month 1 instead of month 13 increases enrollment as of month 24 by 0.73 percentage points, as of month 36 by 0.30 percentage points; and as of month 48 by 0.28 percentage points.[43] These amounts are progressively smaller; implying that those mailed Tickets in month 13 would gradually catch up to those mailed Tickets in month 1.

The second maintained assumption is that the impact of mailing the Ticket on service enrollment for those mailed Tickets in month 13 is always exactly 12 months behind the impact on enrollment for those mailed Tickets in month 1 (hereafter the "total impact not reduced, just delayed" assumption). For instance, the impact of mailing Tickets in month 13 as of month 24, 36 or 48 is exactly the same as the impact of mailing the Ticket in month 1 as of month 12, 24, or 36, respectively. This assumption is illustrated in Exhibit D.1, where we plot total impacts as of months 12, 24, 36 and 48 for Tickets mailed in month 1 (the Xs at 12, 24, 36 and 48 months) and month 13 (the Os at 24, 36, and 48 months). The maintained assumption is that the height of the X at each of months 24, 36 and 48 is the same as the height of the O at month 12, 24, or 36, respectively.

Illustration of how these assumptions are used to produce the projections proceeds inductively. Given the linearity assumption, the solid line segment in Exhibit D.1 connecting the Xs at 1 and 12 months traces the impact of delaying the mailing from month 1 to month 13 over months 1 to 12, and the length of the vertical double arrow at month 12 is the size of the impact at month 12 (1.00, as reported above). That same line segment also represents the total impact over the first 12 months, by definition, because the counterfactual is mailing the Ticket in month 13 or later. Given both assumptions, the first segment of the dashed line (between the Os at months 12 and 24), represents the total impact of mailing the Ticket in month 13 as of month 24. Further, the length of the vertical double arrow between the X and O over the month 24 represents the impact of delaying the mailing from month 1 to month 13 on service enrollment as of that month 24 (0.73 percentage points). Adding the length of this double arrow to the height of the O—the latter being equivalent to the length of the double arrow over 12 months—yields the total impact of mailing the Ticket in month 1 instead of in month 25 or later, as of month 24: 1.73 percentage points (= 1.00 + 0.73).

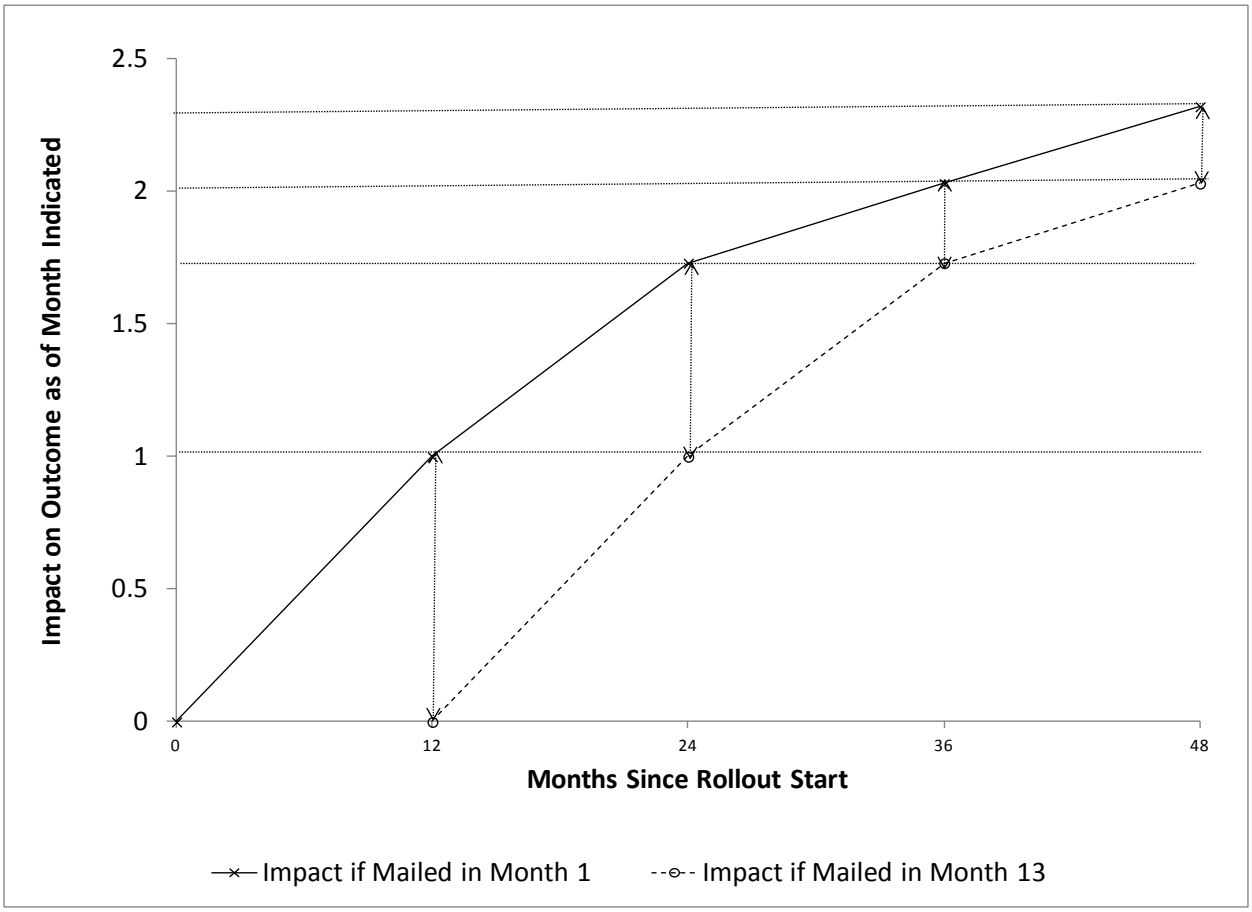
The same logic can be repeated to demonstrate that the sum of the estimated impacts of mailing the Ticket at month 13 instead of month 1 as of months 12, 24 and 36 (illustrated by the lengths of the first three vertical arrows, respectively) is the total impact of mailing the Ticket in month 1 as of month 36—2.03 percentage points (= 1.00 + 0.73 + 0.30). Repeating the logic again yields total impact as of month 48—2.31 percentage points (= 1.00 + 0.73 + 0.30 + 0.28).

If we had instead projected the total impact at 48 months by extrapolating the first segment of the solid line to 48 months—following the simple approach described in the first paragraph of this section—the projected value at month 48 would be 4.0 (four times the projected value at month 12). Our smaller projection at month 48 (2.3)—obtained under a more reasonable set of assumptions coupled with use of the information from estimates for the impacts of delaying the mailing from month 1 to month 13 as of months 24, 36 and 48 as well as month 12—demonstrates that the

[43] The point estimates for the slopes at 24, 36 and 48 months are -0.061, -0.025, and -0.023, respectively.

simple extrapolation would be biased upward at months 24, 36 and 48; although it appears that the total impact increases throughout this period, it does so at a diminishing rate.

Exhibit D.1. Illustrated Computation of the Projection for the Total Impact of Mailing the Ticket on Service Enrollment in Month 48, Phase 2



Note: The diagram plots Phase 2 projections for the total impact of mailing a Ticket in month one on service enrollment at 12, 24, 36 and 48 months under two maintained assumptions: linearity of the impacts of delaying Ticket mailing over the first 13 months of the rollout, and only the timing of the total impact is affected by the length of delay, not its size. The projections are based on IV models for service enrollment with linearity imposed on the MM coefficients. See text for details.

There still might be some upward bias in the projection under the approach used, although we think it is likely to be quite small. If the two maintained assumptions (linearity and total impact not reduced, just delayed) are correct, the projection is unbiased. It seems likely, however, that the second of the two maintained assumptions is somewhat optimistic. Specifically, we would expect impacts for those mailed Tickets in month 13, if anything, to be somewhat smaller than for those mailed Tickets in month 1 because of the passage of 12 months, during which their human capital might well have deteriorated, they might have become better adapted to living on benefits and not engaging in SGA, or they might have managed to find a job or increase their earnings without assistance. If so, then the impact of delay as of 24 months overstates the incremental impact between months 13 and 24 of mailing the Ticket in month 1 instead of month 13, because part of the impact of the delay is the negative effect of waiting on the size of the total impact for those mailed Tickets in month 13 (the O over month 12 in Exhibit D.1 would be lower than drawn, but

the length of the vertical double arrow would be the same). An analogous statement applies to the estimates for the impact of duration to MM over the other 12-month intervals.

It is also important to consider the effect of random variation on the projection. The projection itself is a point estimate. The standard error (SE) for the projection is determined by the four SE for the duration to MM coefficient estimates used in its construction, along with their covariances.[44] For the Phase 2 projection, the SE as of month 48 is 0.76 percentage points, implying that the estimate is significant at the 1 percent level, and that the 95 percent confidence interval ranges from 0.83 to 3.81 percentage points. Thus, the projection of the total impact at 48 months is statistically significant, but the confidence interval around the projection is fairly wide. The confidence intervals around the projections at 12, 24, and 36 months are narrower (0.97 to 1.04, 0.96 to 2.51, and 0.90 to 3.37, respectively), reflecting the fact that fewer estimates are used in their construction.

B. Projections

We applied the conceptual approach described above to all outcome variables in all four samples (Exhibit D.2). Service enrollment projections as of 12, 24, 36 and 48 months appear in the first column. The projected total impact for service enrollment as of 48 months for Phase 1 NY is more than twice the corresponding estimate for Phase 2 (5.7 percentage points), but it is only significant at the 10 percent level because of a relatively large SE—reflecting the relatively short rollout period and relatively small sample. The projected impact for Phase 3 is a much more modest 1.2 percentage points, but it is significant at the 5 percent level despite its smaller size, reflecting that phase's long rollout and large sample. The projection for Phase 1 Except NY, 1.8 percentage points, is comparable in magnitude to the other projections, but not statistically significant.

A final interesting feature of the service enrollment projections is that, in each phase, the point estimates increase with the projection month—reflecting the maintained assumptions and the fact that the restricted IV estimates of all of the coefficients in the duration to MM models for service enrollment are positive. Further, for Phases 2 and 3, the increment to the projection diminishes with each 12-month period, as we would expect. That is not true for the two Phase 1 samples, likely reflecting the relatively large SEs for those estimates.

[44] To produce the standard errors, we used the "three stage least squares" procedure in Stata to jointly estimate the four duration to MM equations for service enrollment for each phase sample, imposing the linearity restriction on the coefficients in each equation and using the IMM for the first-stage instruments.

Exhibit D.2. Projected Total Impacts on Service Enrollment as of 12, 24, 36 and 48 Months, by Phase

Sample and Observation Month	Service Enrollment	TWP Start	TWP Complete	STW	NSTW Months
Phase 1 NY					
12	2.579***	0.037	-0.293	-0.560	-0.090*
z-stat.	3.449	0.038	-0.415	-0.860	-1.762
24	3.668**	-0.013	-0.143	-0.690	-0.181
z-stat.	2.499	-0.006	-0.093	-0.480	-1.086
36	4.367**	-0.668	0.250	0.927	-0.273
z-stat.	2.011	-0.203	0.098	0.390	-0.753
48	5.471*	-2.230	0.044	2.242	-0.286
z-stat.	1.943	-0.482	0.012	0.652	-0.440
Phase 1 Except NY					
12	0.624	-0.679	0.364	0.343	0.034
z-stat.	0.824	-0.711	0.482	0.498	0.624
24	0.876	-1.343	0.793	0.193	0.095
z-stat.	0.592	-0.661	0.487	0.129	0.545
36	1.550	-2.258	1.134	0.996	0.146
z-stat.	0.710	-0.697	0.427	0.406	0.389
48	1.795	-3.556	1.066	1.548	0.212
z-stat.	0.627	-0.782	0.280	0.440	0.319
Phase 2					
12	1.004***	0.260	-0.070	0.001	-0.017
z-stat.	4.895	0.956	-0.326	0.006	-1.114
24	1.736***	0.752	-0.058	0.370	-0.033
z-stat.	4.395	1.285	-0.125	0.892	-0.661
36	2.038***	1.292	-0.016	0.654	-0.009
z-stat.	3.514	1.372	-0.021	0.959	-0.087
48	2.319***	1.749	0.064	1.217	0.062
z-stat.	3.047	1.314	0.058	1.239	0.329
Phase 3					
12	0.791***	-0.008	-0.202	-0.120	-0.005
z-stat.	4.674	-0.037	-1.070	-0.746	-0.353
24	1.028	0.165	0.040	-0.027	-0.035
z-stat.	3.153***	0.337	0.098	-0.076	-0.783
36	1.137	0.636	0.408	0.144	-0.074
z-stat.	2.369	0.808	0.609	0.248	-0.769
48	1.210	0.983	0.899	0.332	-0.140
z-stat.	1.876*	0.888	0.936	0.395	-0.824

Note: "z-stat" is the standard normal test statistic for the hypothesis that the projected impact is zero; *, **, and *** indicate significant at the 0.10, 0.05, and 0.01 levels, respectively.

None of the projections for total impacts on other outcome variables is significant at even the 10 percent level as of any observation point, with the exception of one marginally significant estimate with a sign that is opposite that expected (last four columns of Exhibit D.2). To illustrate, consider the projections for NSTW months. We expect these projections to be positive—especially given the impact on service enrollment—but more are negative than positive, and with one exception all are not close to being statistically significant. The one marginally significant projection has a sign opposite that expected, for month 12 in Phase 1 NY (-0.1 month). The NSTW-months estimates stand in stark contrast to those for service enrollment: uniformly positive point estimates and, apart from the Phase 1 Except NY projections, significant at the 0.10 level or better. These projections reinforce our earlier conclusion that there is no evidence of a substantial impact on any of the outcomes other than service enrollment.

C. Assessment of the Hypothesis That TTW Was Self-Financing by 2007

The fact that we did not find statistically significant impacts on STW or NSTW months does not by itself rule out the possibility that TTW under the initial regulations had impacts on these outcomes that were sufficiently large for the program to be "self-financing"—that is, for savings from a net reduction in benefits to be sufficient to pay for TTW payments to providers and all administrative costs attributed to the program. Thornton (2012) suggests that only a very small impact—an increase of 3,000 or so in the number of all beneficiaries experiencing STW for the first time in each year—might be sufficient for the program to be self-financing. An annual impact on first-time STWs that is as small as 3,000 might correspond to such a small impact on STW as of month 48 for new, young SSD beneficiaries that the evaluation would be unable to differentiate between that impact and no impact at all. This section presents an assessment of whether the evidence from the above analysis allows us to confidently rule out the possibility that TTW was self-financing in 2007—the last full calendar year prior to the change in the regulations.

In what follows, we assume that an annual increase of 3,000 first STW cases is sufficient for TTW to be self-financing. We then develop a minimum value of the percentage impact on STW for new SSD-only beneficiaries that is consistent with an impact of that size for all beneficiaries in 2007. Finally, we assess whether the projections for STW impacts at 48 months provide definitive evidence on whether or not TTW was self-financing prior to the regulatory changes in 2008.

1. Minimum Impact Consistent with Self-Financing

An impact of 3,000 is quite small relative to the number of first-time STW cases actually observed in any recent year. We estimate that this number implies a 5 percent increase in the number of SSD-only beneficiaries having their first STW month in 2007.[45] Under certain strong assumptions a test of the hypothesis that the percentage impact of TTW on STW at 48 months for new, young SSD-only beneficiaries was at least 5 percent could be interpreted as a test of the null hypothesis that TTW was self-financing as of 2007. We consider these assumptions below, and conclude that a 5 percent impact on STW as of month 48 after mailing might be consistent with self-financing as of 2007, but a larger percentage impact would likely be required.

The first assumption required in order to generalize our estimates for young SSD-only beneficiaries to all TTW-eligible beneficiaries is that the percentage impact of TTW on STW for young SSD-only beneficiaries is essentially the same as for other age-program groups. This assumption is at least partially supported by earlier evidence of the percentage impacts of TTW on service enrollment as of the post-rollout year from Stapleton et al. (2008, Exhibit XII.5). They found significant impacts for all age-program groups, ranging from 14 percent to 27 percent; estimates for SSD-only beneficiaries were in the middle of this range (21 percent). If impacts on STW are roughly proportional to impacts on service enrollment, then percentage impacts on STW for SSD-only

[45] The 5 percent value is based on the following data and analysis. Schimmel et al. (2013, Table IV.5) report that, of all SSD and SSI beneficiaries in current pay status in all 12 months of 2006, 73,886 attained STW in 2007, of which almost exactly half (36,768) were SSD-only cases. Hence, if the percentage impact is the same for all program groups, an impact of 3,000 implies an impact of 1,500 for SSD-only cases. The 36,768 SSD-only beneficiaries is an overcount of number of first STW cases since SSD entry because some share of these beneficiaries had attained STW prior to 2006 and then returned to current pay status and remained there throughout 2006 before attaining STW again in 2007. Using the data from Schimmel et al. (2013), we were able to determine that 4,707 of these cases had at least one NSTW month in 2002 through 2005. Netting these out of the SSDI-only cases counted by Schimmel et al. leaves 32,061. The impact of 1,500 for SSD-only cases represents 4.9 percent of this group [(0.049 = 1500/(32061 – 1500)]. We rounded to 5 percent.

beneficiaries ought to be approximately equal to those for all beneficiaries. Of course impacts on STW might not be proportional to impacts on service enrollment, but we do not have evidence on this point.

The second assumption required is that the percentage impact on STW at 48 months after ticket mailing is the same as the percentage impact for all beneficiaries in 2007. This assumption would be reasonable if 2007 had been a "steady state"—that is if, contrary to reality, the number of beneficiaries and the distributions of beneficiary characteristics and other factors that affect the likelihood of STW are not changing from year to year.[46] Under a steady state, the impact of TTW on the percentage of all beneficiaries achieving STW in a single year would be identically equal to the impact on the percentage of new beneficiaries who eventually achieve STW—hereafter the "longitudinal impact". The longitudinal impact is likely very close to the impact of mailing tickets to new beneficiaries on STW as of month 48 after mailing; if mailing a ticket has an impact on STW, it seems very likely that the impact would occur before month 48.

There are two obvious ways that 2007 deviates from a steady state, however, and both imply that the percentage impact on STW at 48 months would have to be larger than 5 percent in order for the percentage impact on *first-ever* STW in 2007 to be 5 percent. We have considered whether there are other causes for deviations from a steady state that might change the percentage in the opposition direction, but have not identified any candidates of note.

The first way that 2007 deviates from a steady state is that the beneficiary population grew very rapidly in the decade leading up to 2007, so the proportion of beneficiaries who had entered in the last decade was much higher in 2007 than it would be in a steady state. Other research implies that a very large percentage of SSD entrants who achieve STW do so in their first 10 years after entry (most in their first five).[47] Hence, the number of first STW cases in 2007 was much larger than it would have been for a steady-state population of the same size. That means that an impact of 3,000 is a larger percentage of the number of cases in a steady-state than of the number of cases in 2007. Because the longitudinal impact is equal to the steady-state impact, it too must be larger than 5 percent.

The second way that 2007 deviates from a steady state is that the TTW rollout had ended only recently, in 2004. If TTW had an impact on STW in 2007, it seems likely that some share of first STWs observed in 2007 represent the impact of providing tickets for the first time to beneficiaries who had been on the rolls for many years—a temporary phenomenon. In a steady-state, the number of first-ever STW cases would be smaller, so an impact of 3,000 cases would be a larger percentage increase. Again this implies that the longitudinal impact would need to exceed 5 percent for STW to be self-financing. We think the effect of this factor on the 2007 percentage is likely to be small relatively to the effect of growth in the beneficiary population, given a total impact of 3,000, because three years after the end of the rollout we would expect most first-ever STW cases to be for relatively new beneficiaries rather than those on the rolls for many years prior to the rollout.

[46] Another way to look at the steady state would be to consider a forward-looking perspective and assess whether impacts in 2007 can be generalized to the current times. Because findings from our impact analysis will not be applicable to TTW under the new regulations, taking this forward looking perspective to the steady state has little relevance here.

[47] Liu and Stapleton (2009) found that, among those who entered SSD in 1996, 4.8 percent had achieved STW by the end of the fifth year, 5.8 percent had achieved STW by the end of the eighth year and only an additional 0.4 percent did so in the next two years.

2. Consistency of the Impact Projections with Self-Financing

Based on Thornton (2012) and the evidence described in the previous section, 5 percent is a candidate for the minimum percentage consistent with self-financing in 2007, but it is likely that the actual minimum percentage impact consistent with self-financing is larger. In this section, we first consider whether the evidence from the projections for the impact at 48 months is consistent with an impact of at least 5 percent. Specifically, we test the following hypothesis: the mailing of tickets to young, new SSD-only beneficiaries increased the number who had attained STW as of month 48 after the mailing by 5 percent versus the alternative hypothesis that the impact was less than 5 percent. We repeat the test for NSTW months, on the assumption that an increase in STW of 5 percent would be sufficient for TTW to be self-financing only if NSTW increases by at least the same relative amount. Finally, we consider how the results would change if the minimum percentage impact consistent with self-financing was larger than 5 percent, as it might well be.

We again focus on Phases 2 and 3 because the power of the STW and NSTW projections for these two phases is much greater than for Phase 1; it is clear that the power of the Phase 1 projections is insufficient to rule out an impact of the required size. We consider the Phase 2 and 3 projections separately, and then, to increase power, we pool the results for the two phases on the assumption that the true relative impacts for the two phases are the same. The pooled projection is the minimum variance projection under the assumption that percentage impacts were the same for Phases 2 and 3.[48] Because of the inequalities in the null and alternative hypotheses, a one-tailed test is appropriate. Results appear in Exhibit D.3. We also show tests for the null hypothesis of "no impact" versus the one-tailed alternative of "positive impact."

For STW, the percentage projections for both phases and the pooled percentage projection are all larger than 5 percent (barely so in Phase 3: 5.4 percent). Hence, in each case we are not able to reject the null hypothesis that the impact is 5 percent points or larger. For the pooled projection, the p-value for the test is 0.74—far above the 0.10 value that is the usual standard for marginal rejection of the null hypothesis. Note, however, that we are also unable to reject the null hypothesis that the impact is zero versus the alternative that it is positive, although the p-value for this test based on the pooled data is much closer to 0.10: 0.14. In short, for this outcome, the evidence is more consistent with an impact of at least 5 percent than with an impact that is zero or negative.

In contrast, for NSTW months, the percentage projections are all *smaller* than 5 percent, and both the Phase 3 and pooled projections are negative (-8.0 percent and -3.4 percent, respectively). We cannot, however, reject the null hypothesis of a 5 percent impact based on the pooled sample (p-value of 0.14). In this case, however, the p-value is much smaller than the p-value for the test of the null hypothesis that the true impact is zero (0.67 percent). That is, for NSTW months the evidence is more consistent with the hypothesis of a zero or negative impact than with an impact of at least 5 percent.

[48] The minimum variance estimate is a weighted mean of the estimates for the two phases where the weights have been chosen to minimize the variance of the estimate. More weight is given to the Phase 3 estimate for each impact because the Phase 3 estimate has lower variance than the Phase 2 estimate.

Exhibit D.3. Projected Relative Impacts on STW and NSTW at 48 Months After Mailing

	Phase 2	Phase 3	Pooled
STW			
Projected relative impacts at 48 months	22.1%	5.4%	11.6%
Standard error of relative impacts	17.8%	13.7%	10.9%
P-value for test of "no impact (or negative impact)" versus "positive impact"	0.108	0.346	0.142
P-value for test of "impact of 5.0% (or more)" versus "impact less than 5.0%"	0.831	0.513	0.729
NSTW			
Projected relative impacts at 48 months	4.1%	-8.0%	-3.4%
Standard error of relative impacts	12.5%	9.7%	7.7%
P-value for test of "no impact (or negative impact)" versus "positive impact"	0.371	0.795	0.673
P-value for test of "impact of 5.0% (or more)" versus "impact less than 5.0%"	0.471	0.090	0.136

Notes: The relative projected impacts were calculated by comparing the projected total impacts on STW and NSTW as of 48 months (see Appendix D, Exhibit D.2) and the estimated means for the corresponding mean in the absence of TTW (counterfactual). For each outcome, the counterfactual mean was estimated by subtracting the weighted mean of the Phase 2 and 3 impact estimates at 48 months from the actual mean for the phase. The p-values are for one-tailed tests, reflecting the inequalities in the hypotheses.

In summary of the analysis to this point, under the strong assumptions discussed above, the statistical power of the projections for STW and NSTW months is insufficient to rule out the possibility that TTW had impacts of at least five percent on each outcome for Phases 2 and 3 pooled. At the same time, the evidence from these projections alone is just as consistent with zero or negative impacts. The pooled projection for STW of 11.6 percent clearly is more consistent with the hypothesis of an impact of at least five percent, but the pooled projection for NSTW months of -3.4 percent is more consistent with a zero or negative impact.

As discussed above, however, there are reasons to believe that the smallest percentage impact estimated for 2007 that is consistent with self-financing is larger than five percent. Because of this, a five percent impact represents a lower bound on the impact necessary for TTW to be self-financing. If we had used a larger value in the tests above, the results would be less favorable to the hypothesis of self-financing. The value used would have to be several times larger for the STW test to lead to rejection of the null-hypothesis at the five-percent significance level when using the pooled data: 30 percent. At the same time, however, the value used would only need to be about twice as large—9 percent—for the NSTW test to lead to rejection of the hypothesis of self-financing at same significance level, using the pooled data.

Thus, although the results overall are consistent with no impact, and we are confident that the estimates are not biased, the statistical power of the methodology is not sufficient to definitively discriminate between "no effect" and "smallest effect consistent with self-financing." Even under the most optimistic assumption that a five percent increase in NSTW months would be sufficient for TTW to be self-financing, the evidence is only marginally consistent with the self-financing hypothesis. If we allow for the more realistic assumptions, then the impact necessary to reach self-financing would need to be higher. If self-financing required at least a nine percent impact on NSTW months—a plausible value—we would have to reject the hypothesis that TTW was self-financing as of 2007.

This page has been left blank for double-sided copying.

APPENDIX E

TABLES WITH IMPACTS AT 12, 24, 36, AND 48 MONTHS FOLLOWING ROLLOUT START: ESTIMATES FROM LINEAR AND INSTRUMENTAL VARIABLES MODELS WITHOUT AND WITH STATE UNEMPLOYMENT MEASURES

This page has been left blank for double-sided copying.

Exhibit E1.a. Linear and IV Model Estimates of Impacts: Service Enrollment

	By Rollout Month 12			By Rollout Month 24			By Rollout Month 36			By Rollout Month 48		
	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat
LM Results for IMM: Impact on Service Enrollment												
Phase 1 NY												
Feb-02	0.0103**	0.004	2.32	0.0060	0.005	1.19	0.0053	0.005	1.03	0.0070	0.005	1.42
May-02	0.0026	0.004	0.60	0.0003	0.005	0.07	-0.0004	0.005	-0.09	-0.0009	0.005	-0.20
July-02	0.0021	0.003	0.69	-0.0013	0.004	-0.36	-0.0029	0.003	-0.85	-0.0018	0.003	-0.58
Aug-02	-0.0043	0.003	-1.53	-0.0010	0.003	-0.31	-0.0004	0.003	-0.14	0.0000	0.003	0.01
Sep-02	-0.0048*	0.003	-1.70	-0.0025	0.003	-0.72	-0.0006	0.003	-0.18	-0.0036	0.003	-1.12
Oct-02	-0.0060	0.003	-2.07	-0.0015	0.004	-0.44	-0.0009	0.003	-0.25	-0.0008	0.0033	-0.24
N	12,023			12,023			12,023			12,023		
Joint sign. test	2.335			0.341			0.285			0.554		
p-value	0.0396			0.888			0.922			0.735		
Test of linear imp	0.337			0.0548			0.286			0.303		
p-value	0.799			0.983			0.836			0.823		
IV Results for MM: Impact on Service Enrollment												
Phase 1 NY												
Feb-02	0.0104**	0.004	2.34	0.0060	0.005	1.19	0.0053	0.005	1.03	0.0071	0.005	1.42
May-02	0.0027	0.004	0.60	0.0004	0.005	0.07	-0.0004	0.005	-0.09	-0.0009	0.005	-0.20
July-02	0.0022	0.003	0.70	-0.0013	0.004	-0.36	-0.0030	0.004	-0.85	-0.0018	0.003	-0.57
Aug-02	-0.0044	0.003	-1.52	-0.0010	0.003	-0.30	-0.0004	0.003	-0.13	0.0001	0.003	0.03
Sep-02	-0.0049*	0.003	-1.68	-0.0025	0.003	-0.72	-0.0006	0.003	-0.18	-0.0036	0.003	-1.12
Oct-02	-0.00605	0.00294	-2.055	-0.00153	0.00359	-0.426	-0.000844	0.00354	-0.239	-0.000756	0.00339	-0.223
N	12,023			12,023			12,023			12,023		
Joint sign. test	11.72			1.713			1.429			2.783		
p-value	0.0388			0.887			0.921			0.733		
Test of linear imp	1.032			0.159			0.851			0.899		
p-value	0.794			0.984			0.837			0.826		

Exhibit E1.a. Linear and IV Model Estimates of Impacts: Service Enrollment

	By Rollout Month 12			By Rollout Month 24			By Rollout Month 36			By Rollout Month 48		
	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat
LM Results for IMM: Impact on Service Enrollment												
Phase 1 Except I												
Feb-02	-0.0004	0.001	-0.34	-0.0004	0.002	-0.21	0.0006	0.002	0.40	-0.0004	0.002	-0.22
Apr-02	0.0018	0.002	1.11	0.0010	0.003	0.40	-0.0000	0.002	-0.02	-0.0006	0.003	-0.20
May-02	-0.0001	0.001	-0.09	-0.0000	0.001	-0.03	0.0010	0.001	0.82	0.0023*	0.001	1.79
Jun-02	-0.0013	0.001	-1.09	-0.0006	0.001	-0.53	-0.0016	0.002	-0.95	-0.0013	0.001	-0.95
N	43,080			43,080			43,080			43,080		
Joint sign. test	0.665			0.120			0.667			1.797		
p-value	0.577			0.948			0.576			0.159		
Test of linear imp	0.806			0.108			0.085			0.251		
p-value	0.374			0.743			0.772			0.618		
IV Results for MM: Impact on Service Enrollment												
Phase 1 Except I												
Feb-02	-0.0009	0.002	-0.52	0.0001	0.003	0.02	0.0036	0.003	1.12	0.0006	0.003	0.21
Apr-02	0.0018	0.002	1.13	0.0010	0.003	0.41	-0.0001	0.002	-0.03	-0.0004	0.003	-0.16
May-02	0.0001	0.001	0.11	-0.0002	0.001	-0.15	-0.0002	0.001	-0.16	0.0018	0.001	1.42
Jun-02	-0.00104	0.00136	-0.763	-0.000873	0.00169	-0.516	-0.00327	0.00187	-1.753	-0.00206	0.00198	-1.037
N	43,043			43,043			43,043			43,043		
Joint sign. test	2.245			0.460			3.716			6.033		
p-value	0.523			0.928			0.294			0.110		
Test of linear imp	0.869			0.101			0.126			0.232		
p-value	0.351			0.751			0.723			0.630		

Exhibit E1.a. Linear and IV Model Estimates of Impacts: Service Enrollment

	By Rollout Month 12			By Rollout Month 24			By Rollout Month 36			By Rollout Month 48		
	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat
LM Results for IMM: Impact on Service Enrollment												
Phase 2												
Nov-02	0.0033***	0.001	3.10	0.0040***	0.001	3.01	0.0019	0.001	1.58	0.0025	0.002	1.60
Jan-03	0.0032**	0.001	2.31	0.0033**	0.001	2.25	0.0029*	0.002	1.78	0.0022	0.001	1.55
Feb-03	0.0026*	0.001	1.90	-0.0004	0.001	-0.30	-0.0016	0.001	-1.09	-0.0012	0.001	-0.97
Mar-03	0.0012	0.002	0.81	0.0006	0.002	0.33	-0.0006	0.001	-0.44	0.0002	0.001	0.19
Apr-03	0.0006	0.001	0.54	0.0018	0.002	1.12	0.0002	0.002	0.10	-0.0011	0.002	-0.54
May-03	-0.0001	0.001	-0.05	-0.0018	0.001	-1.36	-0.0003	0.002	-0.20	-0.0019	0.001	-1.29
Jun-03	-0.0014	0.001	-1.21	-0.0014	0.002	-0.77	0.0000	0.002	0.01	0.0007	0.002	0.37
Jul-03	-0.0042***	0.001	-2.94	-0.0060***	0.001	-4.30	-0.0034***	0.001	-3.00	-0.0031**	0.001	-2.22
Aug-03	-0.0022	0.002	-1.39	0.0005	0.002	0.23	0.0010	0.002	0.64	0.0013	0.001	0.95
Sep-03	-0.0031	0.002	-1.74	-0.0005	0.002	-0.23	-0.0002	0.002	-0.10	0.0003	0.001	0.18
N	77,161			77,161			77,161			77,161		
Joint sign. test	5.191			4.848			2.077			2.050		
p-value	0.000			0.000			0.048			0.051		
Test of linear imp	0.874			2.362			1.933			2.210		
p-value	0.533			0.035			0.083			0.048		
IV Results for MM: Impact on Service Enrollment												
Phase 2												
Nov-02	0.0034***	0.001	3.16	0.0042***	0.001	3.06	0.0020	0.001	1.61	0.0026	0.002	1.62
Jan-03	0.0034**	0.001	2.37	0.0034**	0.001	2.28	0.0030*	0.002	1.80	0.0023	0.001	1.57
Feb-03	0.0028*	0.001	1.94	-0.0005	0.002	-0.29	-0.0016	0.001	-1.10	-0.0013	0.001	-0.97
Mar-03	0.0013	0.002	0.83	0.0006	0.002	0.34	-0.0007	0.002	-0.44	0.0002	0.001	0.19
Apr-03	0.0006	0.001	0.56	0.0019	0.002	1.13	0.0002	0.002	0.10	-0.0011	0.002	-0.54
May-03	-0.0001	0.001	-0.04	-0.0019	0.001	-1.37	-0.0003	0.002	-0.20	-0.0020	0.002	-1.31
Jun-03	-0.0015	0.001	-1.21	-0.0015	0.002	-0.77	0.0000	0.002	0.01	0.0008	0.002	0.38
Jul-03	-0.0044***	0.001	-2.97	-0.0064***	0.001	-4.29	-0.0036***	0.001	-3.01	-0.0033**	0.001	-2.24
Aug-03	-0.0023	0.002	-1.38	0.0005	0.002	0.26	0.0011	0.002	0.65	0.0014	0.001	0.98
Sep-03	-0.0033	0.00189	-1.748	-0.000454	0.00207	-0.219	-0.000173	0.00185	-0.0934	0.000292	0.00156	0.187
N	77,161			77,161			77,161			77,161		
Joint sign. test	49.680			45.060			19.000			18.580		
p-value	0.000			0.000			0.025			0.029		
Test of linear imp	6.504			17.030			13.570			15.550		
p-value	0.482			0.017			0.059			0.030		

Exhibit E1.a. Linear and IV Model Estimates of Impacts: Service Enrollment

	By Rollout Month 12			By Rollout Month 24			By Rollout Month 36			By Rollout Month 48		
	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat
LM Results for IMM: Impact on Service Enrollment												
Phase 3												
Nov-03	0.0040***	0.001	4.40	0.0008	0.001	1.05	0.0020	0.001	1.62	0.0019*	0.001	1.70
Jan-04	0.0016	0.001	1.12	0.0023**	0.001	2.16	0.0010	0.001	0.90	0.0000	0.001	0.04
Feb-04	0.0015**	0.001	2.42	0.0005	0.001	0.64	-0.0010	0.001	-0.81	-0.0009	0.001	-0.80
Mar-04	0.0007	0.001	0.58	-0.0014	0.001	-1.20	-0.0015	0.001	-1.20	-0.0004	0.001	-0.38
Apr-04	-0.0003	0.001	-0.24	-0.0006	0.001	-0.60	-0.0007	0.001	-0.53	-0.0016	0.001	-1.46
May-04	-0.0005	0.001	-0.66	-0.0005	0.001	-0.62	-0.0002	0.001	-0.18	0.0004	0.001	0.33
Jun-04	-0.0012	0.002	-0.80	-0.0009	0.002	-0.53	-0.0011	0.002	-0.57	-0.0004	0.001	-0.28
Jul-04	-0.0005	0.001	-0.45	0.0021*	0.001	1.73	0.0015	0.001	0.99	0.0014	0.001	1.03
Aug-04	-0.0032***	0.001	-3.40	-0.0015	0.001	-1.34	-0.0001	0.001	-0.08	-0.0004	0.001	-0.40
Sep-04	-0.0020	0.001	-2.43	-0.0009	0.001	-0.82	0.0001	0.001	0.09	0.0000	0.001	0.03
N	114,657			114,657			114,657			114,657		
Joint sign. test	5.358			1.555			0.988			1.001		
p-value	0.000			0.153			0.461			0.451		
Test of linear imp	0.868			1.261			0.907			1.184		
p-value	0.538			0.287			0.508			0.328		
IV Results for MM: Impact on Service Enrollment												
Phase 3												
Nov-03	0.0042***	0.001	4.48	0.0009	0.001	1.06	0.0021	0.001	1.63	0.0020*	0.001	1.71
Jan-04	0.0017	0.001	1.14	0.0024**	0.001	2.20	0.001	0.001	0.90	0.0000	0.001	0.03
Feb-04	0.0016**	0.001	2.46	0.0006	0.001	0.64	-0.001	0.001	-0.82	-0.0010	0.001	-0.81
Mar-04	0.0007	0.001	0.60	-0.0015	0.001	-1.22	-0.0015	0.001	-1.22	-0.0004	0.001	-0.38
Apr-04	-0.0003	0.001	-0.24	-0.0006	0.001	-0.61	-0.0007	0.001	-0.53	-0.0017	0.001	-1.45
May-04	-0.0005	0.001	-0.66	-0.0006	0.001	-0.62	-0.0003	0.001	-0.18	0.0004	0.001	0.34
Jun-04	-0.0013	0.002	-0.81	-0.001	0.002	-0.53	-0.0012	0.002	-0.57	-0.0004	0.001	-0.28
Jul-04	-0.0006	0.001	-0.44	0.0023*	0.001	1.74	0.0016	0.002	1.00	0.0015	0.001	1.04
Aug-04	-0.0034***	0.001	-3.42	-0.0016	0.001	-1.35	-0.0001	0.001	-0.08	-0.0005	0.001	-0.40
Sep-04	-0.00216	0.00088	-2.453	-0.000927	0.00114	-0.816	0.000118	0.0013	0.0904	0.0000526	0.00154	0.0341
N	114,657			114,657			114,657			114,657		
Joint sign. test	51.190			14.250			9.163			8.955		
p-value	0.000			0.114			0.422			0.441		
Test of linear imp	6.016			8.953			6.506			8.225		
p-value	0.538			0.256			0.482			0.313		

Exhibit E1.b. Linear and IV Model Estimates of Impacts: TWP Start

	By Rollout Month 12			By Rollout Month 24			By Rollout Month 36			By Rollout Month 48		
	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat
LM Results for IMM: Impact on TWP Start												
Phase 1 NY												
Feb-02	0.0029	0.005	0.55	0.0007	0.007	0.10	0.0010	0.008	0.13	-0.0049	0.008	-0.61
May-02	0.0002	0.005	0.04	0.0049	0.007	0.73	-0.0014	0.007	-0.19	-0.0011	0.008	-0.14
July-02	-0.0055	0.004	-1.51	-0.0061	0.005	-1.29	-0.0061	0.005	-1.12	-0.0035	0.006	-0.60
Aug-02	0.0011	0.004	0.29	-0.0027	0.005	-0.56	0.0001	0.006	0.03	-0.0021	0.006	-0.36
Sep-02	-0.0008	0.004	-0.20	0.0012	0.005	0.24	0.0043	0.006	0.75	0.0108*	0.006	1.75
Oct-02	0.0020	0.004	0.51	0.0020	0.005	0.42	0.0021	0.006	0.38	0.0010	0.006	0.16
N	12,023			12,023			12,023			12,023		
Joint sign. test	0.519			0.466			0.362			0.689		
p-value	0.762			0.802			0.875			0.632		
Test of linear imp	0.652			0.627			0.519			0.717		
p-value	0.582			0.597			0.669			0.542		
IV Results for MM: Impact on TWP Start												
Phase 1 NY												
Feb-02	0.0030	0.005	0.55	0.0007	0.007	0.10	0.0009	0.008	0.12	-0.0049	0.008	-0.62
May-02	0.0002	0.005	0.04	0.0050	0.007	0.73	-0.0014	0.008	-0.19	-0.0012	0.008	-0.14
July-02	-0.0056	0.004	-1.51	-0.0063	0.005	-1.30	-0.0062	0.006	-1.12	-0.0037	0.006	-0.61
Aug-02	0.0012	0.004	0.29	-0.0028	0.005	-0.57	0.0002	0.006	0.03	-0.0022	0.006	-0.36
Sep-02	-0.0008	0.004	-0.20	0.0013	0.005	0.25	0.0044	0.006	0.76	0.0111*	0.006	1.76
Oct-02	0.00209	0.00402	0.519	0.00213	0.00501	0.425	0.00217	0.0057	0.381	0.000923	0.00607	0.152
N	12,023			12,023			12,023			12,023		
Joint sign. test	2.607			2.342			1.817			3.459		
p-value	0.760			0.800			0.874			0.630		
Test of linear imp	1.959			1.886			1.565			2.169		
p-value	0.581			0.596			0.667			0.538		

Exhibit E1.b. Linear and IV Model Estimates of Impacts: TWP Start

	By Rollout Month 12			By Rollout Month 24			By Rollout Month 36			By Rollout Month 48		
	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat
LM Results for IMM: Impact on TWP Start												
Phase 1 Except I												
Feb-02	-0.0022	0.002	-1.20	-0.0022	0.002	-0.93	-0.0016	0.004	-0.45	-0.0036	0.004	-1.03
Apr-02	-0.0010	0.002	-0.65	0.0004	0.002	0.17	-0.0011	0.003	-0.34	0.0008	0.003	0.27
May-02	0.0035**	0.001	2.59	0.0014	0.002	0.78	0.0017	0.002	0.86	0.0013	0.003	0.44
Jun-02	-0.0002	0.002	-0.11	0.0005	0.002	0.26	0.0009	0.002	0.51	0.0015	0.002	0.62
N	43,080			43,080			43,080			43,080		
Joint sign. test	3.165			0.566			0.620			0.759		
p-value	0.032			0.640			0.605			0.522		
Test of linear imp	2.715			0.006			0.194			0.104		
p-value	0.106			0.941			0.662			0.749		
IV Results for MM: Impact on TWP Start												
Phase 1 Except I												
Feb-02	0.0005	0.002	0.20	-0.0003	0.004	-0.06	0.0006	0.006	0.10	-0.0023	0.007	-0.35
Apr-02	-0.0008	0.002	-0.54	0.0005	0.002	0.22	-0.0010	0.003	-0.32	0.0009	0.003	0.29
May-02	0.0024	0.002	1.55	0.0006	0.002	0.30	0.0009	0.002	0.37	0.0008	0.004	0.22
Jun-02	-0.0021	0.00118	-1.778	-0.000864	0.00242	-0.357	-0.000468	0.00282	-0.166	0.000634	0.00348	0.182
N	43,043			43,043			43,043			43,043		
Joint sign. test	6.365			0.239			0.495			0.133		
p-value	0.095			0.971			0.920			0.988		
Test of linear imp	2.049			0.003			0.195			0.090		
p-value	0.152			0.954			0.659			0.764		

Exhibit E1.b. Linear and IV Model Estimates of Impacts: TWP Start

	By Rollout Month 12			By Rollout Month 24			By Rollout Month 36			By Rollout Month 48		
	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat
LM Results for IMM: Impact on TWP Start												
Phase 2												
Nov-02	0.0022	0.002	1.30	0.0021	0.002	1.03	0.0010	0.003	0.39	-0.0002	0.003	-0.05
Jan-03	0.0001	0.001	0.04	0.0015	0.002	0.97	0.0025	0.002	1.32	0.0024	0.002	1.22
Feb-03	-0.0012	0.002	-0.48	0.0009	0.002	0.37	0.0029	0.002	1.38	0.0046*	0.002	1.97
Mar-03	0.0001	0.001	0.08	-0.0019	0.002	-0.86	-0.0024	0.003	-0.86	-0.0010	0.003	-0.32
Apr-03	0.0023	0.003	0.81	0.0025	0.003	0.81	0.0015	0.004	0.43	0.0012	0.003	0.36
May-03	-0.0019	0.002	-0.92	-0.0003	0.002	-0.12	0.0016	0.003	0.59	-0.0005	0.002	-0.22
Jun-03	-0.0001	0.003	-0.02	-0.0002	0.003	-0.06	-0.0028	0.004	-0.79	-0.0046	0.004	-1.29
Jul-03	0.0000	0.002	0.03	-0.0015	0.002	-0.79	-0.0009	0.003	-0.34	-0.0009	0.003	-0.29
Aug-03	0.0001	0.002	0.03	-0.0004	0.002	-0.18	-0.0013	0.003	-0.43	0.0005	0.003	0.18
Sep-03	-0.0016	0.001	-1.07	-0.0028	0.002	-1.36	-0.0022	0.003	-0.82	-0.0016	0.003	-0.59
N	77,161			77,161			77,161			77,161		
Joint sign. test	1.573			1.708			1.386			1.501		
p-value	0.148			0.110			0.218			0.172		
Test of linear imp	0.650			0.453			0.942			1.200		
p-value	0.712			0.863			0.483			0.319		
IV Results for MM: Impact on TWP Start												
Phase 2												
Nov-02	0.0023	0.002	1.31	0.0022	0.002	1.05	0.0011	0.003	0.40	-0.0001	0.003	-0.05
Jan-03	0.0001	0.001	0.05	0.0016	0.002	0.99	0.0027	0.002	1.34	0.0025	0.002	1.24
Feb-03	-0.0012	0.003	-0.48	0.0009	0.003	0.38	0.0031	0.002	1.39	0.0048**	0.002	1.98
Mar-03	0.0001	0.001	0.09	-0.0019	0.002	-0.86	-0.0025	0.003	-0.86	-0.0010	0.003	-0.32
Apr-03	0.0024	0.003	0.81	0.0027	0.003	0.82	0.0016	0.004	0.44	0.0012	0.003	0.37
May-03	-0.0020	0.002	-0.93	-0.0003	0.002	-0.12	0.0016	0.003	0.59	-0.0005	0.002	-0.22
Jun-03	-0.0001	0.003	-0.02	-0.0002	0.003	-0.06	-0.0030	0.004	-0.80	-0.0049	0.004	-1.30
Jul-03	0.0001	0.002	0.03	-0.0016	0.002	-0.80	-0.0009	0.003	-0.34	-0.0009	0.003	-0.29
Aug-03	0.0001	0.002	0.03	-0.0004	0.002	-0.17	-0.0013	0.003	-0.43	0.0006	0.003	0.18
Sep-03	-0.00169	0.00158	-1.069	-0.00291	0.00214	-1.361	-0.00232	0.00284	-0.817	-0.0017	0.00289	-0.589
N	77,161			77,161			77,161			77,161		
Joint sign. test	14.000			14.840			12.280			13.410		
p-value	0.122			0.096			0.198			0.145		
Test of linear imp	4.612			3.219			6.638			8.581		
p-value	0.707			0.864			0.468			0.284		

Exhibit E1.b. Linear and IV Model Estimates of Impacts: TWP Start

	By Rollout Month 12			By Rollout Month 24			By Rollout Month 36			By Rollout Month 48		
	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat
LM Results for IMM: Impact on TWP Start												
Phase 3												
Nov-03	0.0001	0.002	0.04	0.0027*	0.002	1.78	0.0028	0.002	1.60	0.0032	0.002	1.50
Jan-04	-0.0006	0.001	-0.51	0.0016	0.002	0.78	0.0025	0.002	1.12	0.0023	0.002	1.01
Feb-04	0.0002	0.001	0.13	-0.0024	0.002	-1.14	0.0007	0.002	0.34	-0.0008	0.002	-0.35
Mar-04	-0.0005	0.001	-0.52	-0.0027	0.002	-1.61	-0.0021	0.002	-0.90	-0.0026	0.003	-0.89
Apr-04	-0.0012	0.002	-0.77	-0.0015	0.002	-0.84	-0.0026	0.002	-1.22	-0.0037	0.002	-1.66
May-04	0.0037**	0.002	2.32	0.0040	0.002	1.64	0.0031	0.003	1.07	0.0034	0.003	1.07
Jun-04	-0.0006	0.002	-0.40	-0.0008	0.002	-0.41	-0.0014	0.002	-0.65	0.0003	0.002	0.12
Jul-04	-0.0005	0.001	-0.41	0.0001	0.001	0.10	-0.0012	0.002	-0.62	-0.0009	0.002	-0.50
Aug-04	-0.0004	0.001	-0.24	0.0002	0.002	0.12	-0.0006	0.002	-0.27	0.0002	0.002	0.10
Sep-04	-0.0001	0.002	-0.07	-0.0011	0.002	-0.46	-0.0013	0.002	-0.77	-0.0015	0.002	-0.92
N	114,657			114,657			114,657			114,657		
Joint sign. test	0.943			0.760			0.807			0.795		
p-value	0.496			0.653			0.612			0.622		
Test of linear imp	1.161			0.932			0.425			0.728		
p-value	0.341			0.490			0.883			0.649		
IV Results for MM: Impact on TWP Start												
Phase 3												
Nov-03	0.0001	0.002	0.04	0.0028*	0.002	1.80	0.0029	0.002	1.63	0.0034	0.002	1.52
Jan-04	-0.0006	0.001	-0.52	0.0016	0.002	0.78	0.0026	0.002	1.13	0.0025	0.002	1.02
Feb-04	0.0002	0.001	0.13	-0.0026	0.002	-1.15	0.0008	0.002	0.35	-0.0008	0.002	-0.36
Mar-04	-0.0005	0.001	-0.52	-0.0029	0.002	-1.63	-0.0022	0.002	-0.90	-0.0027	0.003	-0.90
Apr-04	-0.0012	0.002	-0.78	-0.0016	0.002	-0.84	-0.0027	0.002	-1.23	-0.0039*	0.002	-1.69
May-04	0.0039**	0.002	2.35	0.0043*	0.003	1.66	0.0034	0.003	1.09	0.0037	0.003	1.09
Jun-04	-0.0007	0.002	-0.40	-0.0008	0.002	-0.41	-0.0015	0.002	-0.66	0.0003	0.002	0.12
Jul-04	-0.0005	0.001	-0.43	0.0001	0.001	0.10	-0.0012	0.002	-0.62	-0.0009	0.002	-0.50
Aug-04	-0.0004	0.002	-0.24	0.0002	0.002	0.12	-0.0006	0.002	-0.27	0.0002	0.002	0.10
Sep-04	-0.000154	0.00208	-0.0739	-0.00118	0.00255	-0.461	-0.00143	0.00186	-0.769	-0.00165	0.00179	-0.925
N	114,657			114,657			114,657			114,657		
Joint sign. test	8.769			7.068			7.369			7.332		
p-value	0.459			0.630			0.599			0.603		
Test of linear imp	8.356			6.713			3.063			5.260		
p-value	0.302			0.459			0.879			0.628		

Exhibit E1.c. Linear and IV Model Estimates of Impacts: TWP Completion

	By Rollout Month 12			By Rollout Month 24			By Rollout Month 36			By Rollout Month 48		
	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat
LM Results for IMM: Impact on TWP Completion												
Phase 1 NY												
Feb-02	-0.0015	0.004	-0.39	0.0005	0.005	0.09	0.0007	0.006	0.11	-0.0046	0.007	-0.67
May-02	0.0030	0.004	0.81	0.0053	0.005	1.00	0.0100	0.007	1.52	0.0092	0.007	1.26
July-02	-0.0021	0.003	-0.78	-0.0048	0.004	-1.30	-0.0068	0.004	-1.52	-0.0044	0.005	-0.87
Aug-02	-0.0022	0.003	-0.80	-0.0030	0.004	-0.77	-0.0044	0.005	-0.97	0.0001	0.005	0.02
Sep-02	0.0007	0.003	0.23	0.0017	0.004	0.43	-0.0027	0.005	-0.60	-0.0018	0.005	-0.34
Oct-02	0.0020	0.003	0.68	0.0003	0.004	0.07	0.0033	0.005	0.68	0.0015	0.005	0.28
N	12,023			12,023			12,023			12,023		
Joint sign. test	0.431			0.571			1.010			0.471		
p-value	0.827			0.723			0.410			0.798		
Test of linear imp	0.505			0.883			1.011			0.724		
p-value	0.679			0.449			0.386			0.537		
IV Results for MM: Impact on TWP Completion												
Phase 1 NY												
Feb-02	-0.0015	0.004	-0.39	0.0005	0.005	0.09	0.0007	0.006	0.11	-0.0046	0.007	-0.67
May-02	0.0031	0.004	0.81	0.0055	0.005	1.00	0.0102	0.007	1.53	0.0094	0.007	1.26
July-02	-0.0021	0.003	-0.79	-0.0050	0.004	-1.31	-0.0070	0.005	-1.54	-0.0046	0.005	-0.88
Aug-02	-0.0023	0.003	-0.81	-0.0031	0.004	-0.78	-0.0047	0.005	-0.99	-0.0000	0.005	-0.00
Sep-02	0.0007	0.003	0.24	0.0018	0.004	0.44	-0.0028	0.005	-0.58	-0.0018	0.005	-0.33
Oct-02	0.0021	0.00307	0.683	0.000342	0.00413	0.0828	0.00345	0.00496	0.696	0.00159	0.00548	0.29
N	12,023			12,023			12,023			12,023		
Joint sign. test	2.166			2.865			5.067			2.365		
p-value	0.826			0.721			0.408			0.797		
Test of linear imp	1.516			2.658			3.047			2.181		
p-value	0.678			0.447			0.384			0.536		

Exhibit E1.c. Linear and IV Model Estimates of Impacts: TWP Completion

	By Rollout Month 12			By Rollout Month 24			By Rollout Month 36			By Rollout Month 48		
	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat
LM Results for IMM: Impact on TWP Completion												
Phase 1 Except I												
Feb-02	0.0003	0.001	0.19	0.0011	0.002	0.52	0.0017	0.003	0.56	0.0009	0.003	0.33
Apr-02	0.0006	0.001	0.56	-0.0008	0.001	-0.57	-0.0014	0.003	-0.46	-0.0020	0.003	-0.64
May-02	-0.0003	0.001	-0.21	0.0006	0.002	0.35	-0.0000	0.003	-0.01	0.0012	0.003	0.45
Jun-02	-0.0006	0.002	-0.35	-0.0009	0.001	-0.66	-0.0004	0.002	-0.18	-0.0001	0.002	-0.05
N	43,080			43,080			43,080			43,080		
Joint sign. test	0.137			0.330			0.141			0.225		
p-value	0.938			0.804			0.935			0.878		
Test of linear imp	0.214			0.671			0.239			0.592		
p-value	0.646			0.416			0.627			0.445		
IV Results for MM: Impact on TWP Completion												
Phase 1 Except I												
Feb-02	0.0009	0.002	0.42	0.0035	0.003	1.36	0.0076*	0.004	1.84	0.0044	0.005	0.87
Apr-02	0.0007	0.001	0.70	-0.0010	0.001	-0.71	-0.0016	0.003	-0.61	-0.0020	0.003	-0.63
May-02	-0.0005	0.002	-0.32	-0.0004	0.002	-0.22	-0.0023	0.003	-0.83	-0.0001	0.003	-0.05
Jun-02	-0.00105	0.00141	-0.746	-0.00216	0.00134	-1.61	-0.00366	0.00189	-1.938	-0.0023	0.00258	-0.888
N	43,043			43,043			43,043			43,043		
Joint sign. test	0.947			3.201			6.314			1.103		
p-value	0.814			0.362			0.097			0.776		
Test of linear imp	0.300			0.933			0.440			0.574		
p-value	0.584			0.334			0.507			0.449		

Exhibit E1.c. Linear and IV Model Estimates of Impacts: TWP Completion

	By Rollout Month 12			By Rollout Month 24			By Rollout Month 36			By Rollout Month 48		
	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat
LM Results for IMM: Impact on TWP Completion												
Phase 2												
Nov-02	-0.0003	0.002	-0.17	0.0007	0.002	0.30	0.0005	0.003	0.20	-0.0002	0.003	-0.08
Jan-03	0.0010	0.001	1.00	0.0011	0.002	0.74	0.0017	0.002	0.74	0.0019	0.002	0.81
Feb-03	-0.0007	0.001	-0.53	-0.0011	0.002	-0.52	-0.0016	0.003	-0.61	-0.0016	0.003	-0.61
Mar-03	-0.0030	0.002	-1.60	-0.0033	0.002	-1.51	-0.0051**	0.002	-2.36	-0.0036*	0.002	-1.76
Apr-03	0.0031**	0.001	2.22	0.0044***	0.001	3.55	0.0069***	0.002	3.62	0.0066***	0.002	3.19
May-03	0.0001	0.001	0.10	-0.0023	0.001	-1.67	-0.0010	0.002	-0.47	-0.0004	0.002	-0.18
Jun-03	-0.0005	0.001	-0.40	0.0012	0.002	0.64	0.0002	0.002	0.10	-0.0013	0.002	-0.56
Jul-03	-0.0012	0.001	-0.91	-0.0023*	0.001	-1.82	-0.0030**	0.001	-2.06	-0.0017	0.002	-0.84
Aug-03	0.0007	0.002	0.43	0.0007	0.002	0.47	0.0012	0.002	0.56	0.0013	0.003	0.50
Sep-03	0.0008	0.002	0.50	0.0009	0.002	0.56	0.0003	0.002	0.15	-0.0010	0.002	-0.42
N	77,161			77,161			77,161			77,161		
Joint sign. test	1.630			5.980			3.144			2.364		
p-value	0.131			0.000			0.004			0.025		
Test of linear imp	1.377			5.242			3.543			2.829		
p-value	0.234			0.000			0.003			0.014		
IV Results for MM: Impact on TWP Completion												
Phase 2												
Nov-02	-0.0004	0.002	-0.18	0.0007	0.002	0.30	0.0005	0.003	0.20	-0.0002	0.003	-0.07
Jan-03	0.0011	0.001	1.01	0.0012	0.002	0.75	0.0017	0.002	0.75	0.0020	0.002	0.82
Feb-03	-0.0007	0.001	-0.53	-0.0011	0.002	-0.53	-0.0017	0.003	-0.61	-0.0017	0.003	-0.61
Mar-03	-0.0031	0.002	-1.61	-0.0035	0.002	-1.53	-0.0053**	0.002	-2.38	-0.0037*	0.002	-1.77
Apr-03	0.0033**	0.001	2.23	0.0046***	0.001	3.56	0.0073***	0.002	3.63	0.0069***	0.002	3.21
May-03	0.0001	0.001	0.09	-0.0025*	0.001	-1.69	-0.0011	0.002	-0.48	-0.0004	0.002	-0.18
Jun-03	-0.0006	0.001	-0.41	0.0013	0.002	0.64	0.0002	0.002	0.10	-0.0014	0.002	-0.57
Jul-03	-0.0013	0.001	-0.93	-0.0024*	0.001	-1.85	-0.0032**	0.002	-2.08	-0.0018	0.002	-0.85
Aug-03	0.0007	0.002	0.44	0.0008	0.002	0.48	0.0012	0.002	0.58	0.0014	0.003	0.51
Sep-03	0.00086	0.00171	0.503	0.000968	0.00172	0.561	0.000278	0.00181	0.153	-0.00107	0.00253	-0.422
N	77,161			77,161			77,161			77,161		
Joint sign. test	14.590			52.160			27.650			21.110		
p-value	0.103			0.000			0.001			0.012		
Test of linear imp	9.741			36.060			24.400			19.750		
p-value	0.204			0.000			0.001			0.006		

Exhibit E1.c. Linear and IV Model Estimates of Impacts: TWP Completion

	By Rollout Month 12			By Rollout Month 24			By Rollout Month 36			By Rollout Month 48		
	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat
LM Results for IMM: Impact on TWP Completion												
Phase 3												
Nov-03	-0.0033**	0.002	-2.14	-0.0019	0.002	-1.21	-0.0011	0.002	-0.69	-0.0002	0.002	-0.14
Jan-04	0.0003	0.001	0.28	0.0040	0.002	1.63	0.0039	0.003	1.43	0.0049	0.003	1.47
Feb-04	0.0009	0.002	0.52	0.0004	0.002	0.23	0.0005	0.002	0.25	0.0007	0.003	0.26
Mar-04	0.0012	0.001	0.95	0.0015	0.001	1.21	0.0008	0.002	0.41	0.0003	0.003	0.12
Apr-04	0.0012	0.001	0.94	-0.0000	0.002	-0.02	0.0002	0.003	0.06	-0.0014	0.003	-0.54
May-04	-0.0005	0.001	-0.47	0.0003	0.002	0.14	0.0022	0.001	1.49	0.0016	0.002	0.97
Jun-04	0.0004	0.001	0.27	-0.0011	0.002	-0.67	-0.0009	0.002	-0.57	-0.0003	0.002	-0.15
Jul-04	-0.0012	0.002	-0.78	-0.0023	0.002	-1.42	-0.0047**	0.002	-2.37	-0.0036*	0.002	-1.95
Aug-04	0.0003	0.002	0.18	-0.0009	0.002	-0.52	-0.0000	0.002	-0.01	-0.0002	0.003	-0.09
Sep-04	0.0007	0.001	0.70	0.0000	0.002	0.00	-0.0008	0.002	-0.44	-0.0018	0.002	-0.86
N	114,657			114,657			114,657			114,657		
Joint sign. test	1.029			2.686			4.700			2.432		
p-value	0.430			0.012			0.000			0.021		
Test of linear imp	1.101			2.261			3.188			1.825		
p-value	0.376			0.043			0.007			0.102		
IV Results for MM: Impact on TWP Completion												
Phase 3												
Nov-03	-0.0035**	0.002	-2.13	-0.0019	0.002	-1.21	-0.0011	0.002	-0.68	-0.0002	0.002	-0.13
Jan-04	0.0003	0.001	0.29	0.0042*	0.003	1.66	0.0042	0.003	1.45	0.0052	0.003	1.49
Feb-04	0.0009	0.002	0.52	0.0004	0.002	0.24	0.0005	0.002	0.26	0.0007	0.003	0.26
Mar-04	0.0013	0.001	0.96	0.0016	0.001	1.21	0.0009	0.002	0.41	0.0003	0.003	0.12
Apr-04	0.0013	0.001	0.94	0.0000	0.002	-0.02	0.0002	0.003	0.06	-0.0015	0.003	-0.54
May-04	-0.0005	0.001	-0.48	0.0003	0.002	0.14	0.0023	0.002	1.50	0.0017	0.002	0.98
Jun-04	0.0004	0.002	0.27	-0.0011	0.002	-0.68	-0.001	0.002	-0.58	-0.0003	0.002	-0.15
Jul-04	-0.0013	0.002	-0.79	-0.0025	0.002	-1.43	-0.0051**	0.002	-2.39	-0.0039**	0.002	-1.97
Aug-04	0.0003	0.002	0.18	-0.001	0.002	-0.52	0.0000	0.002	0.00	-0.0002	0.003	-0.09
Sep-04	0.00078	0.00111	0.705	0.0000169	0.00177	0.00955	-0.00088	0.00201	-0.438	-0.00189	0.00219	-0.86
N	114,657			114,657			114,657			114,657		
Joint sign. test	9.031			24.530			43.020			23.020		
p-value	0.434			0.004			0.000			0.006		
Test of linear imp	7.566			15.890			21.980			12.950		
p-value	0.372			0.026			0.003			0.073		

Exhibit E1.d. Linear and IV Model Estimates of Impacts: STW

	By Rollout Month 12			By Rollout Month 24			By Rollout Month 36			By Rollout Month 48		
	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat
LM Results for IMM: Impact on STW												
Phase 1 NY												
Feb-02	-0.0023	0.003	-0.69	0.0014	0.005	0.29	0.0070	0.006	1.15	0.0018	0.007	0.27
May-02	0.0011	0.003	0.34	-0.0022	0.005	-0.47	0.0076	0.006	1.23	0.0095	0.007	1.39
July-02	-0.0016	0.002	-0.65	-0.0011	0.004	-0.30	-0.0044	0.004	-1.04	-0.0023	0.005	-0.47
Aug-02	0.0016	0.003	0.64	0.0016	0.004	0.44	-0.0055	0.004	-1.29	-0.0009	0.005	-0.19
Sep-02	-0.0031	0.002	-1.28	-0.0021	0.004	-0.59	-0.0044	0.004	-1.00	-0.0073	0.005	-1.50
Oct-02	0.0042	0.003	1.476	0.0024	0.00382	0.619	-0.0002	0.0045	-0.05	-0.0008	0.005	-0.164
N	12,023			12,023			12,023			12,023		
Joint sign. test	0.910			0.226			0.975			0.702		
p-value	0.473			0.951			0.432			0.622		
Test of linear imp	0.707			0.220			0.395			0.622		
p-value	0.548			0.882			0.756			0.600		
IV Results for MM: Impact on STW												
Phase 1 NY												
Feb-02	-0.0023	0.003	-0.69	0.0014	0.005	0.29	0.0070	0.006	1.16	0.0018	0.007	0.28
May-02	0.0011	0.003	0.34	-0.0022	0.005	-0.47	0.0077	0.006	1.23	0.0097	0.007	1.39
July-02	-0.0017	0.003	-0.66	-0.0011	0.004	-0.29	-0.0045	0.004	-1.04	-0.0024	0.005	-0.47
Aug-02	0.0017	0.003	0.63	0.0017	0.004	0.45	-0.0057	0.004	-1.30	-0.0010	0.005	-0.21
Sep-02	-0.0032	0.002	-1.29	-0.0022	0.004	-0.60	-0.0044	0.004	-0.98	-0.0074	0.005	-1.49
Oct-02	0.00434	0.00293	1.483	0.00243	0.00391	0.621	-0.000114	0.00464	-0.0245	-0.000758	0.00515	-0.147
N	12,023			12,023			12,023			12,023		
Joint sign. test	4.570			1.137			4.893			3.522		
p-value	0.471			0.951			0.429			0.620		
Test of linear imp	2.139			0.665			1.194			1.893		
p-value	0.544			0.881			0.754			0.595		

Exhibit E1.d. Linear and IV Model Estimates of Impacts: STW

	By Rollout Month 12			By Rollout Month 24			By Rollout Month 36			By Rollout Month 48		
	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat
LM Results for IMM: Impact on STW												
Phase 1 Except NY												
Feb-02	0.0026	0.002	1.26	0.0000	0.002	0.01	0.0030	0.003	1.09	0.0025	0.003	0.77
Apr-02	-0.0010	0.001	-0.99	-0.0008	0.001	-0.65	-0.0020	0.002	-1.11	-0.0028	0.002	-1.58
May-02	-0.0021***	0.001	-2.74	0.0007	0.002	0.42	0.0002	0.003	0.06	0.0015	0.004	0.41
Jun-02	0.0005	0.002	0.27	0.0001	0.003	0.02	-0.0011	0.002	-0.48	-0.0012	0.002	-0.47
N	43,080			43,080			43,080			43,080		
Joint sign. test	3.456			0.250			0.770			1.116		
p-value	0.023			0.861			0.516			0.351		
Test of linear imp	0.074			0.616			1.283			2.555		
p-value	0.787			0.436			0.263			0.116		
IV Results for MM: Impact on STW												
Phase 1 Except NY												
Feb-02	0.0012	0.002	0.65	0.0015	0.003	0.45	0.0063	0.005	1.29	0.0074	0.005	1.53
Apr-02	-0.0005	0.001	-0.49	-0.0004	0.001	-0.33	-0.0020	0.002	-1.19	-0.0028*	0.002	-1.80
May-02	-0.0016	0.001	-1.07	0.0001	0.003	0.05	-0.0011	0.004	-0.32	-0.0005	0.004	-0.12
Jun-02	0.00086	0.00115	0.75	-0.00128	0.00163	-0.782	-0.00319	0.00224	-1.423	-0.00405	0.00208	-1.945
N	43,043			43,043			43,043			43,043		
Joint sign. test	2.020			1.208			2.696			5.165		
p-value	0.568			0.751			0.441			0.160		
Test of linear imp	0.003			0.495			1.413			2.963		
p-value	0.956			0.482			0.235			0.085		

Exhibit E1.d. Linear and IV Model Estimates of Impacts: STW

	By Rollout Month 12			By Rollout Month 24			By Rollout Month 36			By Rollout Month 48		
	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat
LM Results for IMM: Impact on STW												
Phase 2												
Nov-02	-0.0016	0.001	-1.44	0.0007	0.001	0.49	0.0005	0.002	0.29	0.0013	0.002	0.56
Jan-03	0.0009	0.001	1.02	0.0002	0.002	0.14	-0.0018	0.002	-0.85	-0.0003	0.002	-0.11
Feb-03	0.0012	0.001	0.94	0.0024	0.001	1.67	0.0025	0.002	1.44	0.0026	0.002	1.41
Mar-03	-0.0007	0.002	-0.45	-0.0007	0.002	-0.37	-0.0001	0.002	-0.06	-0.0000	0.003	-0.00
Apr-03	-0.0005	0.001	-0.50	0.0001	0.002	0.05	0.0015	0.002	0.64	0.0034	0.002	1.41
May-03	0.0008	0.001	0.64	0.0009	0.001	0.64	0.0015	0.002	0.74	0.0003	0.003	0.12
Jun-03	0.0004	0.001	0.42	0.0012	0.002	0.77	-0.0004	0.002	-0.20	-0.0023	0.003	-0.93
Jul-03	0.0019	0.001	1.27	0.0001	0.001	0.04	0.0016	0.002	1.06	0.0017	0.002	1.04
Aug-03	-0.0006	0.001	-0.59	-0.0022	0.002	-1.25	-0.0021	0.003	-0.83	-0.0025	0.003	-0.76
Sep-03	-0.0017	0.002	-1.04	-0.0027	0.002	-1.66	-0.0032	0.002	-1.977	-0.0042	0.002	-2.37
N	77,161			77,161			77,161			77,161		
Joint sign. test	1.064			1.389			1.330			1.775		
p-value	0.404			0.217			0.244			0.095		
Test of linear imp	1.170			1.503			0.576			0.729		
p-value	0.335			0.186			0.773			0.648		
IV Results for MM: Impact on STW												
Phase 2												
Nov-02	-0.0016	0.001	-1.46	0.0007	0.001	0.50	0.0005	0.002	0.30	0.0013	0.002	0.57
Jan-03	0.0009	0.001	1.03	0.0003	0.002	0.15	-0.0018	0.002	-0.85	-0.0003	0.003	-0.10
Feb-03	0.0012	0.001	0.95	0.0025*	0.001	1.70	0.0026	0.002	1.46	0.0027	0.002	1.43
Mar-03	-0.0007	0.002	-0.45	-0.0008	0.002	-0.37	-0.0001	0.002	-0.06	0.0000	0.003	0.00
Apr-03	-0.0005	0.001	-0.50	0.0001	0.002	0.06	0.0016	0.002	0.65	0.0036	0.002	1.43
May-03	0.0008	0.001	0.65	0.0010	0.001	0.65	0.0015	0.002	0.75	0.0003	0.003	0.12
Jun-03	0.0005	0.001	0.43	0.0013	0.002	0.78	-0.0004	0.002	-0.20	-0.0024	0.003	-0.93
Jul-03	0.0020	0.002	1.28	0.0001	0.002	0.04	0.0017	0.002	1.07	0.0019	0.002	1.05
Aug-03	-0.0007	0.001	-0.60	-0.0023	0.002	-1.26	-0.0022	0.003	-0.83	-0.0026	0.003	-0.77
Sep-03	-0.00181	0.00172	-1.055	-0.00281	0.00168	-1.673	-0.00336	0.00168	-2.004	-0.00449	0.00187	-2.399
N	77,161			77,161			77,161			77,161		
Joint sign. test	9.900			13.190			12.560			16.290		
p-value	0.359			0.154			0.184			0.061		
Test of linear imp	8.536			11.16			4.141			5.192		
p-value	0.288			0.132			0.763			0.636		

Exhibit E1.d. Linear and IV Model Estimates of Impacts: STW

	By Rollout Month 12			By Rollout Month 24			By Rollout Month 36			By Rollout Month 48		
	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat
LM Results for IMM: Impact on STW												
Phase 3												
Nov-03	-0.0023***	0.001	-3.32	-0.0020**	0.001	-2.03	-0.0017	0.001	-1.21	-0.0016	0.001	-1.33
Jan-04	0.0019	0.001	1.39	0.0024*	0.001	1.70	0.0041*	0.002	1.71	0.0030	0.003	1.03
Feb-04	-0.0006	0.001	-0.83	0.0006	0.001	0.65	-0.0004	0.002	-0.28	-0.0001	0.002	-0.04
Mar-04	0.0005	0.001	0.54	0.0003	0.002	0.21	-0.0008	0.002	-0.52	-0.0008	0.002	-0.41
Apr-04	-0.0002	0.001	-0.27	-0.0008	0.002	-0.43	-0.0002	0.002	-0.09	-0.0005	0.003	-0.17
May-04	-0.0004	0.001	-0.35	0.0018	0.001	1.30	0.0028	0.002	1.54	0.0054***	0.002	2.97
Jun-04	0.0022***	0.001	3.02	0.0013	0.001	1.15	0.0010	0.001	0.73	0.0003	0.001	0.25
Jul-04	-0.0013	0.001	-1.29	-0.0030**	0.001	-2.41	-0.0036**	0.001	-2.42	-0.0045**	0.002	-2.52
Aug-04	-0.0005	0.001	-0.60	0.0003	0.002	0.17	-0.0007	0.002	-0.31	0.0004	0.003	0.14
Sep-04	0.0007	0.001	0.81	-0.0009	0.001	-1.10	-0.0004	0.002	-0.28	-0.0016	0.001	-1.08
N	114,657			114,657			114,657			114,657		
Joint sign. test	4.769			2.297			2.452			3.606		
p-value	0.000			0.029			0.021			0.001		
Test of linear imp	4.421			2.465			1.768			3.059		
p-value	0.001			0.029			0.113			0.009		
IV Results for MM: Impact on STW												
Phase 3												
Nov-03	-0.0024***	0.001	-3.39	-0.0021**	0.001	-2.06	-0.0018	0.001	-1.23	-0.0017	0.001	-1.35
Jan-04	0.002	0.001	1.41	0.0025*	0.001	1.72	0.0043*	0.002	1.73	0.0032	0.003	1.04
Feb-04	-0.0007	0.001	-0.83	0.0006	0.001	0.66	-0.0005	0.002	-0.28	-0.0001	0.002	-0.04
Mar-04	0.0005	0.001	0.55	0.0003	0.002	0.21	-0.0009	0.002	-0.53	-0.0009	0.002	-0.41
Apr-04	-0.0002	0.001	-0.28	-0.0008	0.002	-0.44	-0.0002	0.002	-0.09	-0.0005	0.003	-0.17
May-04	-0.0004	0.001	-0.35	0.002	0.002	1.31	0.003	0.002	1.55	0.0057***	0.002	3.00
Jun-04	0.0024***	0.001	3.01	0.0014	0.001	1.15	0.001	0.001	0.73	0.0004	0.001	0.25
Jul-04	-0.0014	0.001	-1.30	-0.0032**	0.001	-2.43	-0.0038**	0.002	-2.43	-0.0048**	0.002	-2.55
Aug-04	-0.0006	0.001	-0.61	0.0003	0.002	0.17	-0.0007	0.002	-0.31	0.0004	0.003	0.14
Sep-04	0.00077	0.00095	0.812	-0.00101	0.00092	-1.102	-0.000452	0.00163	-0.277	-0.00169	0.00157	-1.079
N	114,657			114,657			114,657			114,657		
Joint sign. test	42.030			20.960			22.300			33.000		
p-value	0.000			0.013			0.008			0.000		
Test of linear imp	29.430			17.540			12.880			22.280		
p-value	0.000			0.014			0.075			0.002		

Exhibit E1.e. Linear and IV Model Estimates of Impacts: NSTW Months

	By Rollout Month 12			By Rollout Month 24			By Rollout Month 36			By Rollout Month 48		
	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat
LM Results for IMM: Impact on NSTW Months												
Phase 1 NY												
Feb-02	-0.0517**	0.026	-2.03	-0.0233	0.063	-0.37	-0.0066	0.107	-0.06	0.0235	0.159	0.15
May-02	0.0074	0.025	0.29	-0.0283	0.059	-0.48	-0.0259	0.102	-0.25	-0.0174	0.154	-0.11
July-02	0.0165	0.020	0.82	0.0175	0.045	0.39	-0.0035	0.077	-0.05	-0.0133	0.113	-0.12
Aug-02	-0.0046	0.020	-0.22	-0.0064	0.048	-0.13	-0.0306	0.081	-0.38	-0.0141	0.118	-0.12
Sep-02	0.0020	0.019	0.10	-0.0116	0.046	-0.25	0.0020	0.080	0.03	-0.0175	0.118	-0.15
Oct-02	0.0305	0.022	1.38	0.0521	0.053	0.99	0.0646	0.089	0.73	0.0388	0.128	0.30
N	12,023			12,023			12,023			12,023		
Joint sign. test	1.085			0.262			0.126			0.0296		
p-value	0.366			0.934			0.987			1.000		
Test of linear imp	0.701			0.124			0.0337			0.00435		
p-value	0.552			0.946			0.992			1.000		
IV Results for MM: Impact on NSTW Months												
Phase 1 NY												
Feb-02	-0.0520**	0.026	-2.03	-0.0236	0.063	-0.38	-0.0068	0.108	-0.06	0.0235	0.160	0.15
May-02	0.0075	0.026	0.29	-0.0289	0.060	-0.48	-0.0266	0.104	-0.25	-0.0178	0.157	-0.11
July-02	0.0166	0.020	0.81	0.0179	0.046	0.39	-0.0036	0.079	-0.05	-0.0135	0.116	-0.12
Aug-02	-0.0051	0.021	-0.24	-0.0064	0.049	-0.13	-0.0311	0.083	-0.38	-0.0141	0.121	-0.12
Sep-02	0.0018	0.019	0.09	-0.0123	0.047	-0.26	0.0019	0.082	0.02	-0.0180	0.121	-0.15
Oct-02	0.0311	0.0226	1.373	0.0533	0.0539	0.99	0.0663	0.0907	0.731	0.04	0.131	0.306
N	12,023			12,023			12,023			12,023		
Joint sign. test	5.449			1.314			0.633			0.149		
p-value	0.364			0.933			0.986			1.000		
Test of linear imp	2.082			0.374			0.101			0.013		
p-value	0.556			0.946			0.992			1.000		

Exhibit E1.e. Linear and IV Model Estimates of Impacts: NSTW Months

	By Rollout Month 12			By Rollout Month 24			By Rollout Month 36			By Rollout Month 48		
	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat
LM Results for IMM: Impact on NSTW Months												
Phase 1 Except I												
Feb-02	-0.0029	0.023	-0.12	-0.0136	0.054	-0.25	-0.0215	0.071	-0.30	-0.0253	0.084	-0.30
Apr-02	0.0146	0.014	1.03	0.0347	0.027	1.31	0.0300	0.036	0.83	0.0271	0.050	0.54
May-02	-0.0069	0.008	-0.90	-0.0100	0.018	-0.56	0.0082	0.035	0.24	0.0241	0.056	0.43
Jun-02	-0.0048	0.014	-0.34	-0.0111	0.038	-0.29	-0.0167	0.059	-0.29	-0.0259	0.079	-0.33
N	43,080			43,080			43,080			43,080		
Joint sign. test	0.655			1.150			0.622			0.619		
p-value	0.584			0.338			0.604			0.606		
Test of linear imp	0.825			0.970			0.210			0.0377		
p-value	0.368			0.329			0.648			0.847		
IV Results for MM: Impact on NSTW Months												
Phase 1 Except I												
Feb-02	-0.0115	0.019	-0.62	-0.0274	0.035	-0.79	-0.0353	0.053	-0.67	-0.0204	0.076	-0.27
Apr-02	0.0181	0.015	1.18	0.0436	0.029	1.51	0.0411	0.038	1.09	0.0393	0.051	0.78
May-02	-0.0036	0.010	-0.38	-0.0048	0.029	-0.16	0.0135	0.054	0.25	0.0220	0.082	0.27
Jun-02	-0.00292	0.0107	-0.274	-0.0113	0.0187	-0.606	-0.0193	0.0348	-0.554	-0.0409	0.0525	-0.779
N	43,043			43,043			43,043			43,043		
Joint sign. test	1.661			3.391			2.709			2.776		
p-value	0.646			0.335			0.439			0.427		
Test of linear imp	1.150			1.508			0.458			0.113		
p-value	0.284			0.219			0.499			0.736		

Exhibit E1.e. Linear and IV Model Estimates of Impacts: NSTW Months

	By Rollout Month 12			By Rollout Month 24			By Rollout Month 36			By Rollout Month 48		
	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat
LM Results for IMM: Impact on NSTW Months												
Phase 2												
Nov-02	-0.0242**	0.010	-2.38	-0.0408**	0.019	-2.12	-0.0318	0.026	-1.20	-0.0247	0.044	-0.56
Jan-03	0.0064	0.008	0.81	0.0215	0.024	0.91	0.0223	0.045	0.50	0.0230	0.066	0.35
Feb-03	0.0051	0.007	0.72	0.0132	0.020	0.66	0.0361	0.032	1.14	0.0716	0.045	1.58
Mar-03	-0.0026	0.013	-0.20	-0.0170	0.034	-0.51	-0.0294	0.054	-0.54	-0.0309	0.080	-0.39
Apr-03	-0.0039	0.010	-0.38	-0.0054	0.026	-0.21	0.0020	0.043	0.05	0.0179	0.056	0.32
May-03	0.0020	0.011	0.19	0.0162	0.024	0.68	0.0449	0.039	1.14	0.0566	0.057	0.99
Jun-03	0.0169*	0.010	1.76	0.0366	0.024	1.53	0.0421	0.043	0.99	0.0312	0.067	0.47
Jul-03	0.0018	0.013	0.14	-0.0040	0.026	-0.15	-0.0244	0.035	-0.71	-0.0351	0.039	-0.90
Aug-03	0.0028	0.010	0.28	-0.0080	0.027	-0.30	-0.0339	0.045	-0.75	-0.0765	0.065	-1.17
Sep-03	-0.0044	0.013	-0.34	-0.0123	0.029	-0.43	-0.0279	0.045	-0.62	-0.0332	0.059	-0.57
N	77,161			77,161			77,161			77,161		
Joint sign. test	0.988			1.324			1.984			1.984		
p-value	0.461			0.247			0.060			0.060		
Test of linear imp	1.133			1.671			1.685			1.255		
p-value	0.357			0.136			0.133			0.290		
IV Results for MM: Impact on NSTW Months												
Phase 2												
Nov-02	-0.0251**	0.010	-2.42	-0.0422**	0.020	-2.15	-0.0327	0.027	-1.21	-0.0251	0.045	-0.56
Jan-03	0.0067	0.008	0.81	0.0224	0.024	0.92	0.0234	0.046	0.51	0.0244	0.069	0.36
Feb-03	0.0053	0.007	0.72	0.0137	0.021	0.66	0.0377	0.033	1.16	0.0749	0.047	1.60
Mar-03	-0.0027	0.013	-0.21	-0.0177	0.035	-0.51	-0.0305	0.057	-0.54	-0.0320	0.083	-0.39
Apr-03	-0.0041	0.011	-0.38	-0.0056	0.027	-0.21	0.0022	0.045	0.05	0.0191	0.059	0.32
May-03	0.0021	0.011	0.19	0.0171	0.025	0.69	0.0474	0.041	1.15	0.0597	0.060	1.00
Jun-03	0.0178*	0.010	1.78	0.0386	0.025	1.54	0.0443	0.045	1.00	0.0328	0.070	0.47
Jul-03	0.0019	0.014	0.14	-0.0044	0.028	-0.16	-0.0262	0.036	-0.72	-0.0376	0.041	-0.91
Aug-03	0.0029	0.011	0.28	-0.0086	0.028	-0.31	-0.0360	0.047	-0.76	-0.0810	0.068	-1.18
Sep-03	-0.00478	0.0137	-0.35	-0.0132	0.0303	-0.436	-0.0296	0.0473	-0.626	-0.0351	0.0621	-0.565
N	77,161			77,161			77,161			77,161		
Joint sign. test	9.129			12.120			17.780			17.850		
p-value	0.425			0.207			0.038			0.037		
Test of linear imp	8.056			11.870			12.000			9.035		
p-value	0.328			0.105			0.101			0.250		

Exhibit E1.e. Linear and IV Model Estimates of Impacts: NSTW Months

	By Rollout Month 12			By Rollout Month 24			By Rollout Month 36			By Rollout Month 48		
	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat
LM Results for IMM: Impact on NSTW Months												
Phase 3												
Nov-03	-0.0032	0.009	-0.36	-0.0228	0.021	-1.09	-0.0386	0.036	-1.07	-0.0656	0.053	-1.25
Jan-04	0.0164**	0.008	2.07	0.0363**	0.016	2.25	0.0666**	0.032	2.07	0.0861	0.052	1.65
Feb-04	-0.0082	0.006	-1.39	-0.0237	0.017	-1.41	-0.0401	0.029	-1.36	-0.0404	0.048	-0.85
Mar-04	-0.0032	0.009	-0.35	-0.0027	0.020	-0.13	-0.0026	0.035	-0.08	-0.0316	0.046	-0.69
Apr-04	-0.0014	0.007	-0.20	-0.0110	0.020	-0.54	-0.0230	0.040	-0.57	-0.0268	0.066	-0.41
May-04	-0.0251***	0.009	-2.83	-0.0358	0.022	-1.61	-0.0333	0.035	-0.96	-0.0092	0.050	-0.18
Jun-04	0.0110	0.011	1.00	0.0241	0.026	0.92	0.0358	0.040	0.90	0.0589	0.047	1.25
Jul-04	0.0027	0.005	0.58	0.0072	0.014	0.50	-0.0011	0.021	-0.05	-0.0258	0.030	-0.86
Aug-04	0.0106	0.008	1.32	0.0184	0.019	0.97	0.0234	0.036	0.65	0.0261	0.053	0.50
Sep-04	0.0004	0.006	0.06	0.0100	0.016	0.64	0.0130	0.022	0.59	0.0284	0.033	0.86
N	114,657			114,657			114,657			114,657		
Joint sign. test	4.814			2.027			1.679			2.372		
p-value	0.000			0.054			0.117			0.025		
Test of linear imp	4.551			2.351			1.787			1.949		
p-value	0.000			0.036			0.109			0.080		
IV Results for MM: Impact on NSTW Months												
Phase 3												
Nov-03	-0.0034	0.009	-0.37	-0.0241	0.022	-1.11	-0.0407	0.037	-1.09	-0.0691	0.055	-1.26
Jan-04	0.0173**	0.008	2.10	0.0383**	0.017	2.28	0.0704**	0.034	2.09	0.0910*	0.055	1.66
Feb-04	-0.0087	0.006	-1.41	-0.0252	0.018	-1.43	-0.0425	0.031	-1.38	-0.043	0.05	-0.86
Mar-04	-0.0034	0.009	-0.36	-0.003	0.021	-0.14	-0.0029	0.036	-0.08	-0.0338	0.048	-0.71
Apr-04	-0.0015	0.007	-0.21	-0.0118	0.021	-0.55	-0.0246	0.042	-0.58	-0.0287	0.069	-0.41
May-04	-0.0269***	0.009	-2.85	-0.0383	0.024	-1.63	-0.0356	0.037	-0.97	-0.0098	0.053	-0.19
Jun-04	0.0118	0.012	1.01	0.0258	0.028	0.93	0.0383	0.042	0.91	0.063	0.050	1.26
Jul-04	0.003	0.005	0.59	0.0077	0.015	0.51	-0.0012	0.022	-0.06	-0.028	0.032	-0.88
Aug-04	0.0114	0.009	1.34	0.0197	0.020	0.98	0.0250	0.038	0.65	0.0279	0.056	0.50
Sep-04	0.000414	0.00624	0.0663	0.0107	0.0166	0.646	0.0139	0.0234	0.594	0.0305	0.0352	0.866
N	114,657			114,657			114,657			114,657		
Joint sign. test	46.270			19.170			15.720			22.880		
p-value	0.000			0.024			0.073			0.006		
Test of linear imp	33.890			17.200			13.100			14.430		
p-value	0.000			0.016			0.070			0.044		

Exhibit E2.a. Impacts on Service Enrollment: Linear and IV Models with Controls for State Level Unemployment

	By Rollout Month 12			By Rollout Month 24			By Rollout Month 36			By Rollout Month 48		
	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat
LM Results for IMM: Impact on Service Enrollment												
Phase 1 Except NY												
Feb-02	-0.0009	0.002	-0.51	0.0001	0.003	0.02	0.0036	0.003	1.10	0.0007	0.003	0.21
Apr-02	0.0018	0.002	1.12	0.0010	0.003	0.40	-0.0001	0.002	-0.04	-0.0004	0.003	-0.15
May-02	0.0001	0.001	0.11	-0.0002	0.001	-0.15	-0.0002	0.001	-0.17	0.0018	0.001	1.40
Jun-02	-0.0010	0.001	-0.75	-0.0009	0.002	-0.51	-0.0033	0.002	-1.73	-0.0020	0.002	-1.03
N	43,043			43,043			43,043			43,043		
Joint sign. test	0.731			0.150			1.211			1.965		
p-value	0.538			0.929			0.316			0.132		
Test of linear impact	0.850			0.0986			0.124			0.226		
p-value	0.361			0.755			0.726			0.637		
IV Results for MM: Impact on Service Enrollment												
Phase 1 Except NY												
Feb-02	-0.0009	0.002	-0.52	0.0001	0.003	0.02	0.0036	0.003	1.12	0.0006	0.003	0.21
Apr-02	0.0018	0.002	1.13	0.0010	0.003	0.41	-0.0001	0.002	-0.03	-0.0004	0.003	-0.16
May-02	0.0001	0.001	0.11	-0.0002	0.001	-0.15	-0.0002	0.001	-0.16	0.0018	0.001	1.42
Jun-02	-0.0010	0.001	-0.76	-0.0009	0.002	-0.52	-0.0033	0.002	-1.75	-0.0021	0.002	-1.04
N	43,043			43,043			43,043			43,043		
Joint sign. test	2.245			0.460			3.716			6.033		
p-value	0.523			0.928			0.294			0.110		
Test of linear impact	0.869			0.101			0.126			0.232		
p-value	0.351			0.751			0.723			0.630		

Exhibit E2.a. Impacts on Service Enrollment: Linear and IV Models with Controls for State Level Unemployment

	By Rollout Month 12			By Rollout Month 24			By Rollout Month 36			By Rollout Month 48		
	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat
LM Results for IMM: Impact on Service Enrollment												
Phase 2												
Nov-02	0.0033**	0.001	2.54	0.0034*	0.002	1.80	0.0014	0.002	0.76	0.0029	0.002	1.48
Jan-03	0.0033*	0.002	1.89	0.0040**	0.002	2.24	0.0031*	0.002	1.68	0.0028	0.002	1.67
Feb-03	0.0026	0.002	1.44	0.0009	0.001	0.60	-0.0010	0.002	-0.67	-0.0007	0.001	-0.46
Mar-03	0.0012	0.002	0.69	0.0025	0.002	1.41	0.0002	0.002	0.14	0.0008	0.002	0.48
Apr-03	0.0006	0.001	0.43	0.0037*	0.002	1.80	0.0011	0.002	0.47	-0.0007	0.002	-0.30
May-03	-0.0001	0.001	-0.06	-0.0004	0.001	-0.28	0.0004	0.002	0.22	-0.0018	0.002	-1.04
Jun-03	-0.0015	0.001	-1.16	-0.0010	0.002	-0.53	0.0003	0.002	0.13	0.0005	0.002	0.26
Jul-03	-0.0042***	0.002	-2.73	-0.0070***	0.002	-3.81	-0.0037**	0.001	-2.65	-0.0036**	0.002	-2.05
Aug-03	-0.0022	0.003	-0.81	-0.0021	0.003	-0.76	-0.0000	0.002	-0.02	0.0004	0.002	0.21
Sep-03	-0.0031	0.003	-1.17	-0.0039	0.003	-1.52	-0.0017	0.003	-0.56	-0.0007	0.003	-0.23
N	77,128			77,128			77,128			77,128		
Joint sign. test	1.690			2.772			1.494			2.075		
p-value	0.116			0.0101			0.176			0.0494		
Test of linear impact	0.698			2.259			1.739			1.858		
p-value	0.673			0.0441			0.121			0.0963		
IV Results for MM: Impact on Service Enrollment												
Phase 2												
Nov-02	0.0036***	0.001	2.62	0.0037*	0.002	1.89	0.0015	0.002	0.80	0.0031	0.002	1.52
Jan-03	0.0035*	0.002	1.96	0.0044**	0.002	2.31	0.0033*	0.002	1.70	0.0030*	0.002	1.68
Feb-03	0.0029	0.002	1.49	0.0011	0.002	0.72	-0.0010	0.002	-0.62	-0.0007	0.002	-0.42
Mar-03	0.0014	0.002	0.75	0.0028	0.002	1.51	0.0003	0.002	0.17	0.0008	0.002	0.50
Apr-03	0.0007	0.001	0.47	0.0040*	0.002	1.83	0.0012	0.003	0.48	-0.0007	0.002	-0.29
May-03	-0.0001	0.001	-0.05	-0.0004	0.002	-0.26	0.0005	0.002	0.23	-0.0019	0.002	-1.05
Jun-03	-0.0016	0.001	-1.20	-0.0011	0.002	-0.56	0.0003	0.002	0.13	0.0006	0.002	0.25
Jul-03	-0.0045***	0.002	-2.78	-0.0076***	0.002	-3.83	-0.0040***	0.001	-2.67	-0.0039**	0.002	-2.06
Aug-03	-0.0024	0.003	-0.85	-0.0024	0.003	-0.81	-0.0001	0.002	-0.05	0.0004	0.002	0.20
Sep-03	-0.0035	0.003	-1.20	-0.0044	0.003	-1.58	-0.0019	0.003	-0.57	-0.0008	0.003	-0.23
N	77,128			77,128			77,128			77,128		
Joint sign. test	16.41			26.23			13.70			19.57		
p-value	0.0588			0.00187			0.134			0.0207		
Test of linear impact	5.302			16.18			12.16			13.03		
p-value	0.623			0.0236			0.0953			0.0715		

Exhibit E2.a. Impacts on Service Enrollment: Linear and IV Models with Controls for State Level Unemployment

	By Rollout Month 12			By Rollout Month 24			By Rollout Month 36			By Rollout Month 48		
	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat
LM Results for IMM: Impact on Service Enrollment												
Phase 3												
Nov-03	0.0046***	0.001	4.46	0.0015	0.001	1.60	0.0025	0.002	1.59	0.0026*	0.001	1.80
Jan-04	0.0020	0.001	1.44	0.0028**	0.001	2.44	0.0014	0.001	1.12	0.0006	0.001	0.49
Feb-04	0.0016**	0.001	2.05	0.0010	0.001	1.19	-0.0005	0.001	-0.43	-0.0004	0.001	-0.37
Mar-04	0.0008	0.001	0.65	-0.0012	0.001	-1.03	-0.0014	0.001	-1.15	-0.0003	0.001	-0.27
Apr-04	-0.0003	0.001	-0.27	-0.0006	0.001	-0.59	-0.0006	0.001	-0.51	-0.0015	0.001	-1.42
May-04	-0.0005	0.001	-0.67	-0.0007	0.001	-0.76	-0.0003	0.001	-0.23	0.0003	0.001	0.24
Jun-04	-0.0013	0.001	-0.92	-0.0012	0.002	-0.69	-0.0014	0.002	-0.68	-0.0007	0.001	-0.47
Jul-04	-0.0008	0.001	-0.59	0.0016	0.001	1.24	0.0010	0.002	0.66	0.0009	0.001	0.63
Aug-04	-0.0035***	0.001	-3.27	-0.0019	0.001	-1.57	-0.0004	0.001	-0.37	-0.0009	0.001	-0.78
Sep-04	-0.0025	0.001	-2.62	-0.0014	0.001	-1.14	-0.0002	0.001	-0.18	-0.0005	0.002	-0.30
N	114,377			114,377			114,377			114,377		
Joint sign. test	3.889			1.793			1.190			1.020		
p-value	0.000827			0.0924			0.321			0.437		
Test of linear impact	0.864			1.260			0.834			1.135		
p-value	0.541			0.289			0.564			0.356		
IV Results for MM: Impact on Service Enrollment												
Phase 3												
Nov-03	0.0050***	0.001	4.65	0.0017	0.001	1.63	0.0026	0.002	1.60	0.0028*	0.002	1.81
Jan-04	0.0022	0.001	1.52	0.0030**	0.001	2.50	0.0015	0.001	1.14	0.0006	0.001	0.51
Feb-04	0.0018**	0.001	2.13	0.0011	0.001	1.21	-0.0005	0.001	-0.42	-0.0004	0.001	-0.36
Mar-04	0.0009	0.001	0.68	-0.0013	0.001	-1.03	-0.0014	0.001	-1.16	-0.0003	0.001	-0.26
Apr-04	-0.0003	0.001	-0.26	-0.0006	0.001	-0.58	-0.0007	0.001	-0.51	-0.0016	0.001	-1.41
May-04	-0.0005	0.001	-0.68	-0.0007	0.001	-0.76	-0.0003	0.001	-0.23	0.0003	0.001	0.24
Jun-04	-0.0014	0.002	-0.94	-0.0013	0.002	-0.70	-0.0015	0.002	-0.69	-0.0007	0.002	-0.48
Jul-04	-0.0009	0.001	-0.62	0.0017	0.001	1.22	0.0011	0.002	0.66	0.0010	0.002	0.63
Aug-04	-0.0039***	0.001	-3.29	-0.0021	0.001	-1.59	-0.0005	0.001	-0.38	-0.0010	0.001	-0.79
Sep-04	-0.0027	0.001	-2.70	-0.0015	0.001	-1.15	-0.0003	0.002	-0.18	-0.0005	0.002	-0.31
N	114,377			114,377			114,377			114,377		
Joint sign. test	38.02			16.35			11.16			9.183		
p-value	1.73e-05			0.0599			0.265			0.421		
Test of linear impact	6.020			8.945			5.991			7.892		
p-value	0.537			0.257			0.541			0.342		

Exhibit E2.b. Impacts on TWP Start: Linear and IV Models with Controls for State Level Unemployment

	By Rollout Month 12			By Rollout Month 24			By Rollout Month 36			By Rollout Month 48		
	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat
LM Results for IMM: Impact on TWP Start												
Phase 1 Except NY												
Feb-02	0.0005	0.003	0.20	-0.0003	0.004	-0.06	0.0006	0.006	0.10	-0.0023	0.007	-0.34
Apr-02	-0.0008	0.002	-0.53	0.0005	0.002	0.21	-0.0010	0.003	-0.31	0.0009	0.003	0.29
May-02	0.0024	0.002	1.53	0.0006	0.002	0.29	0.0009	0.002	0.36	0.0008	0.004	0.21
Jun-02	-0.0021	0.001	-1.76	-0.0009	0.002	-0.35	-0.0005	0.003	-0.17	0.0006	0.003	0.18
N	43,043			43,043			43,043			43,043		
Joint sign. test	2.074			0.0778			0.161			0.0432		
p-value	0.116			0.972			0.922			0.988		
Test of linear impact	1.994			0.00326			0.189			0.0883		
p-value	0.164			0.955			0.665			0.768		
IV Results for MM: Impact on TWP Start												
Phase 1 Except NY												
Feb-02	0.0005	0.002	0.20	-0.0003	0.004	-0.06	0.0006	0.006	0.10	-0.0023	0.007	-0.35
Apr-02	-0.0008	0.002	-0.54	0.0005	0.002	0.22	-0.0010	0.003	-0.32	0.0009	0.003	0.29
May-02	0.0024	0.002	1.55	0.0006	0.002	0.30	0.0009	0.002	0.37	0.0008	0.004	0.22
Jun-02	-0.0021	0.001	-1.78	-0.0009	0.002	-0.36	-0.0005	0.003	-0.17	0.0006	0.003	0.18
N	43,043			43,043			43,043			43,043		
Joint sign. test	6.365			0.239			0.495			0.133		
p-value	0.0951			0.971			0.920			0.988		
Test of linear impact	2.049			0.00332			0.195			0.0899		
p-value	0.152			0.954			0.659			0.764		

Exhibit E2.b. Impacts on TWP Start: Linear and IV Models with Controls for State Level Unemployment

	By Rollout Month 12			By Rollout Month 24			By Rollout Month 36			By Rollout Month 48		
	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat
LM Results for IMM: Impact on TWP Start												
Phase 2												
Nov-02	0.0007	0.002	0.34	0.0001	0.002	0.05	0.0002	0.003	0.06	-0.0006	0.003	-0.19
Jan-03	-0.0020	0.002	-1.03	-0.0010	0.002	-0.55	0.0008	0.002	0.33	0.0010	0.002	0.44
Feb-03	-0.0033	0.003	-1.18	-0.0017	0.003	-0.65	0.0008	0.003	0.30	0.0028	0.003	0.89
Mar-03	-0.0019	0.002	-1.00	-0.0044*	0.002	-1.93	-0.0047	0.003	-1.52	-0.0030	0.004	-0.84
Apr-03	0.0007	0.003	0.29	0.0006	0.003	0.21	-0.0004	0.004	-0.11	-0.0007	0.003	-0.23
May-03	-0.0025	0.002	-1.15	-0.0010	0.002	-0.41	0.0005	0.003	0.20	-0.0016	0.002	-0.70
Jun-03	0.0006	0.003	0.24	0.0007	0.003	0.22	-0.0025	0.004	-0.64	-0.0045	0.004	-1.09
Jul-03	0.0021	0.002	1.04	0.0011	0.002	0.48	0.0010	0.003	0.33	0.0007	0.003	0.22
Aug-03	0.0034	0.003	1.36	0.0037	0.003	1.30	0.0022	0.004	0.56	0.0037	0.004	0.94
Sep-03	0.0021	0.002	0.98	0.0019	0.003	0.66	0.0020	0.004	0.56	0.0023	0.004	0.63
N	77,128			77,128			77,128			77,128		
Joint sign. test	0.783			0.645			0.767			1.022		
p-value	0.632			0.753			0.647			0.435		
Test of linear impact	0.915			0.600			0.963			1.202		
p-value	0.503			0.753			0.468			0.319		
IV Results for MM: Impact on TWP Start												
Phase 2												
Nov-02	0.0006	0.002	0.30	0.0000	0.002	0.01	0.0001	0.003	0.05	-0.0007	0.003	-0.20
Jan-03	-0.0022	0.002	-1.08	-0.0012	0.002	-0.60	0.0008	0.003	0.30	0.0010	0.002	0.41
Feb-03	-0.0035	0.003	-1.22	-0.0019	0.003	-0.69	0.0008	0.003	0.27	0.0028	0.003	0.86
Mar-03	-0.0021	0.002	-1.04	-0.0047*	0.002	-1.95	-0.0050	0.003	-1.52	-0.0032	0.004	-0.85
Apr-03	0.0007	0.003	0.26	0.0006	0.003	0.19	-0.0005	0.004	-0.12	-0.0008	0.003	-0.25
May-03	-0.0027	0.002	-1.18	-0.0011	0.003	-0.43	0.0005	0.003	0.19	-0.0017	0.002	-0.72
Jun-03	0.0007	0.003	0.25	0.0007	0.003	0.23	-0.0026	0.004	-0.64	-0.0047	0.004	-1.10
Jul-03	0.0023	0.002	1.09	0.0012	0.002	0.52	0.0011	0.003	0.34	0.0008	0.004	0.23
Aug-03	0.0037	0.003	1.39	0.0041	0.003	1.32	0.0024	0.004	0.58	0.0040	0.004	0.95
Sep-03	0.0024	0.002	1.03	0.0022	0.003	0.69	0.0022	0.004	0.58	0.0025	0.004	0.65
N	77,128			77,128			77,128			77,128		
Joint sign. test	7.287			6.039			6.900			9.314		
p-value	0.607			0.736			0.648			0.409		
Test of linear impact	6.652			4.219			6.724			8.490		
p-value	0.466			0.754			0.458			0.291		

Exhibit E2.b. Impacts on TWP Start: Linear and IV Models with Controls for State Level Unemployment

	By Rollout Month 12			By Rollout Month 24			By Rollout Month 36			By Rollout Month 48		
	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat
LM Results for IMM: Impact on TWP Start												
Phase 3												
Nov-03	-0.0006	0.002	-0.33	0.0024	0.002	1.32	0.0031	0.002	1.66	0.0038*	0.002	1.74
Jan-04	-0.0010	0.001	-0.82	0.0013	0.002	0.65	0.0025	0.002	1.13	0.0025	0.002	1.10
Feb-04	-0.0002	0.001	-0.11	-0.0027	0.002	-1.19	0.0006	0.002	0.26	-0.0009	0.002	-0.39
Mar-04	-0.0008	0.001	-0.72	-0.0029*	0.002	-1.68	-0.0021	0.002	-0.93	-0.0026	0.003	-0.90
Apr-04	-0.0011	0.002	-0.74	-0.0017	0.002	-0.89	-0.0027	0.002	-1.29	-0.0038*	0.002	-1.72
May-04	0.0037**	0.002	2.28	0.0040	0.002	1.61	0.0030	0.003	1.02	0.0033	0.003	1.02
Jun-04	-0.0004	0.002	-0.24	-0.0006	0.002	-0.28	-0.0013	0.002	-0.60	0.0003	0.002	0.14
Jul-04	-0.0001	0.001	-0.06	0.0005	0.001	0.35	-0.0010	0.002	-0.51	-0.0008	0.002	-0.43
Aug-04	0.0001	0.001	0.04	0.0005	0.002	0.25	-0.0006	0.002	-0.30	-0.0000	0.002	-0.00
Sep-04	0.0004	0.002	0.21	-0.0008	0.003	-0.30	-0.0015	0.002	-0.73	-0.0019	0.002	-1.03
N	114,377			114,377			114,377			114,377		
Joint sign. test	0.874			0.875			0.602			1.078		
p-value	0.554			0.553			0.789			0.395		
Test of linear impact	1.046			1.036			0.462			0.938		
p-value	0.412			0.418			0.858			0.486		
IV Results for MM: Impact on TWP Start												
Phase 3												
Nov-03	-0.0007	0.002	-0.34	0.0025	0.002	1.31	0.0034*	0.002	1.69	0.0040*	0.002	1.77
Jan-04	-0.0011	0.001	-0.84	0.0013	0.002	0.65	0.0027	0.002	1.14	0.0027	0.002	1.11
Feb-04	-0.0002	0.001	-0.12	-0.0029	0.002	-1.20	0.0007	0.002	0.27	-0.0009	0.002	-0.38
Mar-04	-0.0008	0.001	-0.74	-0.0031*	0.002	-1.70	-0.0022	0.002	-0.94	-0.0027	0.003	-0.91
Apr-04	-0.0012	0.002	-0.75	-0.0018	0.002	-0.90	-0.0029	0.002	-1.30	-0.0041*	0.002	-1.75
May-04	0.0039**	0.002	2.31	0.0043	0.003	1.63	0.0032	0.003	1.04	0.0036	0.003	1.04
Jun-04	-0.0004	0.002	-0.24	-0.0006	0.002	-0.28	-0.0014	0.002	-0.61	0.0003	0.002	0.13
Jul-04	-0.0001	0.001	-0.06	0.0005	0.002	0.35	-0.0011	0.002	-0.52	-0.0008	0.002	-0.44
Aug-04	0.0001	0.001	0.04	0.0005	0.002	0.25	-0.0007	0.002	-0.32	-0.0000	0.002	-0.01
Sep-04	0.0005	0.002	0.21	-0.0008	0.003	-0.30	-0.0016	0.002	-0.74	-0.0021	0.002	-1.04
N	114,377			114,377			114,377			114,377		
Joint sign. test	8.167			8.165			5.453			9.863		
p-value	0.517			0.518			0.793			0.362		
Test of linear impact	7.557			7.459			3.327			6.765		
p-value	0.373			0.383			0.853			0.454		

Exhibit E2.c. Impacts on TWP Completion: Linear and IV Models with Controls for State Level Unemployment

	By Rollout Month 12			By Rollout Month 24			By Rollout Month 36			By Rollout Month 48		
	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat
LM Results for IMM: Impact on TWP Completion												
Phase 1 Except NY												
Feb-02	0.0009	0.002	0.41	0.0035	0.003	1.34	0.0076*	0.004	1.81	0.0044	0.005	0.86
Apr-02	0.0007	0.001	0.69	-0.0010	0.001	-0.70	-0.0016	0.003	-0.60	-0.0020	0.003	-0.62
May-02	-0.0005	0.002	-0.31	-0.0004	0.002	-0.23	-0.0023	0.003	-0.83	-0.0001	0.003	-0.05
Jun-02	-0.0010	0.00141	-0.738	-0.0022	0.001	-1.60	-0.0036	0.002	-1.92	-0.0023	0.003	-0.88
N	43,043			43,043			43,043			43,043		
Joint sign. test	0.309			1.044			2.059			0.359		
p-value	0.819			0.381			0.118			0.782		
Test of linear impact	0.291			0.918			0.436			0.563		
p-value	0.592			0.343			0.512			0.457		
IV Results for MM: Impact on TWP Completion												
Phase 1 Except NY												
Feb-02	0.0009	0.002	0.42	0.0035	0.003	1.36	0.0076*	0.004	1.84	0.0044	0.005	0.87
Apr-02	0.0007	0.001	0.70	-0.0010	0.001	-0.71	-0.0016	0.003	-0.61	-0.0020	0.003	-0.63
May-02	-0.0005	0.002	-0.32	-0.0004	0.002	-0.22	-0.0023	0.003	-0.83	-0.0001	0.003	-0.05
Jun-02	-0.0011	0.00141	-0.75	-0.0022	0.001	-1.61	-0.0037	0.002	-1.94	-0.0023	0.003	-0.89
N	43,043			43,043			43,043			43,043		
Joint sign. test	0.947			3.201			6.314			1.103		
p-value	0.814			0.362			0.0973			0.776		
Test of linear impact	0.300			0.933			0.440			0.574		
p-value	0.584			0.334			0.507			0.449		

Exhibit E2.c. Impacts on TWP Completion: Linear and IV Models with Controls for State Level Unemployment

	By Rollout Month 12			By Rollout Month 24			By Rollout Month 36			By Rollout Month 48		
	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat
LM Results for IMM: Impact on TWP Completion												
Phase 2												
Nov-02	-0.0011	0.002	-0.50	0.0004	0.003	0.12	0.0006	0.003	0.17	0.0005	0.004	0.11
Jan-03	-0.0002	0.002	-0.15	0.0001	0.002	0.03	0.0004	0.003	0.15	0.0022	0.003	0.80
Feb-03	-0.0021	0.002	-1.15	-0.0025	0.002	-1.06	-0.0035	0.003	-1.19	-0.0016	0.004	-0.46
Mar-03	-0.0045**	0.002	-2.14	-0.0049*	0.003	-1.89	-0.0075***	0.002	-3.27	-0.0038*	0.002	-1.80
Apr-03	0.0019	0.002	1.01	0.0030*	0.002	1.93	0.0047*	0.002	1.95	0.0062**	0.003	2.31
May-03	-0.0005	0.001	-0.42	-0.0032**	0.001	-2.16	-0.0025	0.002	-1.14	-0.0009	0.002	-0.38
Jun-03	-0.0002	0.001	-0.16	0.0013	0.002	0.65	0.0001	0.002	0.05	-0.0017	0.003	-0.58
Jul-03	0.0001	0.002	0.06	-0.0010	0.002	-0.51	-0.0014	0.002	-0.63	-0.0019	0.003	-0.67
Aug-03	0.0031	0.003	1.19	0.0031	0.003	1.13	0.0045	0.003	1.46	0.0015	0.004	0.41
Sep-03	0.0036	0.003	1.35	0.0039	0.002	1.59	0.0047	0.003	1.82	-0.0005	0.003	-0.15
N	77,128			77,128			77,128			77,128		
Joint sign. test	1.788			7.412			4.646			2.070		
p-value	0.0936			8.01e-07			0.000163			0.0499		
Test of linear impact	1.632			5.343			5.059			2.167		
p-value	0.148			0.000125			0.000208			0.0528		
IV Results for MM: Impact on TWP Completion												
Phase 2												
Nov-02	-0.0012	0.002	-0.53	0.0003	0.003	0.10	0.0005	0.003	0.14	0.0005	0.004	0.12
Jan-03	-0.0004	0.002	-0.20	-0.0000	0.002	-0.02	0.0003	0.003	0.10	0.0023	0.003	0.79
Feb-03	-0.0023	0.002	-1.16	-0.0027	0.002	-1.09	-0.0038	0.003	-1.22	-0.0017	0.004	-0.45
Mar-03	-0.0048**	0.002	-2.14	-0.0052*	0.003	-1.91	-0.0079***	0.002	-3.31	-0.0039*	0.002	-1.76
Apr-03	0.0019	0.002	0.97	0.0030*	0.002	1.85	0.0048*	0.003	1.90	0.0065**	0.003	2.31
May-03	-0.0006	0.001	-0.44	-0.0034**	0.002	-2.21	-0.0027	0.002	-1.18	-0.0009	0.002	-0.39
Jun-03	-0.0002	0.001	-0.15	0.0014	0.002	0.66	0.0001	0.003	0.05	-0.0018	0.003	-0.59
Jul-03	0.0002	0.002	0.10	-0.0010	0.002	-0.47	-0.0014	0.002	-0.59	-0.0020	0.003	-0.67
Aug-03	0.0034	0.003	1.21	0.0034	0.003	1.14	0.0050	0.003	1.49	0.0016	0.004	0.40
Sep-03	0.0040	0.003	1.35	0.0043	0.003	1.59	0.0052	0.003	1.84	-0.0005	0.003	-0.16
N	77,128			77,128			77,128			77,128		
Joint sign. test	15.81			62.89			40.31			18.62		
p-value	0.0708			3.71e-10			6.67e-06			0.0286		
Test of linear impact	11.44			36.81			34.43			15.23		
p-value	0.120			5.10e-06			1.43e-05			0.0331		

Exhibit E2.c. Impacts on TWP Completion: Linear and IV Models with Controls for State Level Unemployment

	By Rollout Month 12			By Rollout Month 24			By Rollout Month 36			By Rollout Month 48		
	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat
LM Results for IMM: Impact on TWP Completion												
Phase 3												
Nov-03	-0.0031*	0.002	-1.69	-0.0017	0.002	-0.94	0.0003	0.002	0.15	0.0009	0.002	0.45
Jan-04	0.0004	0.001	0.36	0.0040	0.003	1.57	0.0046	0.003	1.55	0.0053	0.004	1.51
Feb-04	0.0009	0.002	0.57	0.0002	0.002	0.14	0.0008	0.002	0.42	0.0008	0.003	0.32
Mar-04	0.0012	0.001	0.96	0.0015	0.001	1.24	0.0011	0.002	0.57	0.0004	0.003	0.15
Apr-04	0.0012	0.001	0.91	-0.0000	0.002	-0.01	0.0002	0.003	0.09	-0.0015	0.003	-0.57
May-04	-0.0005	0.001	-0.47	0.0003	0.002	0.14	0.0021	0.001	1.47	0.0016	0.002	0.95
Jun-04	0.0004	0.001	0.27	-0.0010	0.002	-0.60	-0.0013	0.002	-0.73	-0.0004	0.002	-0.21
Jul-04	-0.0013	0.002	-0.81	-0.0023	0.002	-1.38	-0.0053***	0.002	-2.79	-0.0039**	0.002	-2.09
Aug-04	0.0001	0.001	0.09	-0.0010	0.002	-0.53	-0.0008	0.002	-0.34	-0.0008	0.003	-0.28
Sep-04	0.0005	0.001	0.43	0.0000	0.002	0.01	-0.0018	0.002	-0.88	-0.0024	0.002	-0.99
N	114,377			114,377			114,377			114,377		
Joint sign. test	0.997			2.405			4.396			1.729		
p-value	0.454			0.0234			0.000277			0.106		
Test of linear impact	1.082			2.200			3.046			1.570		
p-value	0.389			0.0496			0.00936			0.166		
IV Results for MM: Impact on TWP Completion												
Phase 3												
Nov-03	-0.0032*	0.002	-1.66	-0.0017	0.002	-0.90	0.0004	0.002	0.22	0.0010	0.002	0.50
Jan-04	0.0005	0.001	0.36	0.0042	0.003	1.60	0.0049	0.003	1.58	0.0057	0.004	1.54
Feb-04	0.0010	0.002	0.57	0.0003	0.002	0.16	0.0009	0.002	0.46	0.0009	0.003	0.34
Mar-04	0.0013	0.001	0.97	0.0016	0.001	1.26	0.0012	0.002	0.59	0.0004	0.003	0.17
Apr-04	0.0013	0.001	0.91	-0.0000	0.002	-0.01	0.0003	0.003	0.09	-0.0015	0.003	-0.58
May-04	-0.0006	0.001	-0.48	0.0003	0.002	0.14	0.0023	0.002	1.48	0.0017	0.002	0.96
Jun-04	0.0004	0.002	0.27	-0.0011	0.002	-0.62	-0.0014	0.002	-0.76	-0.0004	0.002	-0.23
Jul-04	-0.0014	0.002	-0.81	-0.0025	0.002	-1.41	-0.0057***	0.002	-2.85	-0.0043**	0.002	-2.14
Aug-04	0.0001	0.002	0.09	-0.0011	0.002	-0.55	-0.0009	0.003	-0.36	-0.0009	0.003	-0.30
Sep-04	0.0006	0.001	0.43	0.0000	0.002	-0.01	-0.0020	0.002	-0.90	-0.0027	0.003	-1.01
N	114,377			114,377			114,377			114,377		
Joint sign. test	8.817			22.05			40.41			16.06		
p-value	0.454			0.00871			6.41e-06			0.0656		
Test of linear impact	7.462			15.54			21.14			11.23		
p-value	0.382			0.0297			0.00356			0.129		

Exhibit E2.d. Impacts on STW: Linear and IV Models with Controls for State Level Unemployment

	By Rollout Month 12			By Rollout Month 24			By Rollout Month 36			By Rollout Month 48		
	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat
LM Results for IMM: Impact on STW												
Phase 1 Except NY												
Feb-02	0.0012	0.002	0.65	0.0015	0.003	0.44	0.0063	0.005	1.28	0.0073	0.005	1.51
Apr-02	-0.0005	0.001	-0.49	-0.0004	0.001	-0.33	-0.0020	0.002	-1.18	-0.0028*	0.002	-1.79
May-02	-0.0016	0.001	-1.06	0.0001	0.003	0.04	-0.0012	0.004	-0.32	-0.0005	0.004	-0.12
Jun-02	0.0009	0.001	0.74	-0.0013	0.002	-0.774	-0.0032	0.002	-1.41	-0.0040	0.002	-1.93
N	43,043			43,043			43,043			43,043		
Joint sign. test	0.657			0.394			0.879			1.684		
p-value	0.582			0.758			0.458			0.183		
Test of linear impact	0.00272			0.488			1.393			2.919		
p-value	0.959			0.488			0.244			0.0939		
IV Results for MM: Impact on STW												
Phase 1 Except NY												
Feb-02	0.0012	0.002	0.65	0.0015	0.003	0.45	0.0063	0.005	1.29	0.0074	0.005	1.53
Apr-02	-0.0005	0.001	-0.49	-0.0004	0.001	-0.33	-0.0020	0.002	-1.19	-0.0028*	0.002	-1.80
May-02	-0.0016	0.001	-1.07	0.0001	0.003	0.05	-0.0011	0.004	-0.32	-0.0005	0.004	-0.12
Jun-02	0.0009	0.001	0.75	-0.0013	0.002	-0.78	-0.0032	0.002	-1.42	-0.0041	0.002	-1.95
N	43,043			43,043			43,043			43,043		
Joint sign. test	2.020			1.208			2.696			5.165		
p-value	0.568			0.751			0.441			0.160		
Test of linear impact	0.00304			0.495			1.413			2.963		
p-value	0.956			0.482			0.235			0.0852		

Exhibit E2.d. Impacts on STW: Linear and IV Models with Controls for State Level Unemployment

	By Rollout Month 12			By Rollout Month 24			By Rollout Month 36			By Rollout Month 48		
	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat
LM Results for IMM: Impact on STW												
Phase 2												
Nov-02	-0.0010	0.001	-0.69	0.0020	0.002	1.00	0.0013	0.002	0.65	0.0024	0.003	0.90
Jan-03	0.0014	0.001	1.01	0.0002	0.002	0.11	-0.0021	0.002	-0.93	0.0000	0.003	0.00
Feb-03	0.0016	0.001	1.06	0.0016	0.002	0.95	0.0016	0.002	0.76	0.0023	0.003	0.93
Mar-03	-0.0004	0.002	-0.24	-0.0021	0.002	-0.95	-0.0015	0.002	-0.81	-0.0007	0.002	-0.30
Apr-03	-0.0004	0.001	-0.33	-0.0015	0.002	-0.72	-0.0000	0.003	-0.01	0.0024	0.003	0.82
May-03	0.0007	0.001	0.47	-0.0005	0.001	-0.35	0.0002	0.002	0.12	-0.0007	0.003	-0.27
Jun-03	0.0001	0.001	0.12	0.0005	0.002	0.29	-0.0009	0.002	-0.40	-0.0030	0.003	-1.00
Jul-03	0.0014	0.001	0.97	0.0004	0.002	0.25	0.0022	0.002	1.29	0.0017	0.002	0.76
Aug-03	-0.0012	0.002	-0.76	-0.0006	0.003	-0.22	-0.0003	0.003	-0.10	-0.0017	0.003	-0.52
Sep-03	-0.0022	0.003	-0.80	0.0000	0.002	-0.02	-0.0005	0.002	-0.23	-0.0028	0.003	-1.01
N	77,128			77,128			77,128			77,128		
Joint sign. test	1.037			0.895			1.063			1.090		
p-value	0.424			0.537			0.405			0.387		
Test of linear impact	1.289			1.142			0.815			0.710		
p-value	0.275			0.352			0.579			0.664		
IV Results for MM: Impact on STW												
Phase 2												
Nov-02	-0.0010	0.001	-0.67	0.0021	0.002	1.00	0.0013	0.002	0.64	0.0026	0.003	0.91
Jan-03	0.0015	0.001	1.02	0.0002	0.002	0.11	-0.0022	0.002	-0.94	0.0001	0.003	0.02
Feb-03	0.0017	0.002	1.07	0.0017	0.002	0.94	0.0017	0.002	0.75	0.0025	0.003	0.94
Mar-03	-0.0004	0.002	-0.22	-0.0022	0.002	-0.95	-0.0016	0.002	-0.82	-0.0007	0.002	-0.28
Apr-03	-0.0004	0.001	-0.30	-0.0016	0.002	-0.72	-0.0000	0.003	-0.02	0.0026	0.003	0.83
May-03	0.0007	0.001	0.49	-0.0006	0.002	-0.36	0.0002	0.002	0.12	-0.0007	0.003	-0.26
Jun-03	0.0001	0.001	0.11	0.0005	0.002	0.29	-0.0009	0.002	-0.40	-0.0031	0.003	-1.01
Jul-03	0.0014	0.002	0.95	0.0004	0.002	0.25	0.0023	0.002	1.30	0.0018	0.002	0.74
Aug-03	-0.0013	0.002	-0.77	-0.0006	0.003	-0.21	-0.0003	0.003	-0.09	-0.0019	0.004	-0.54
Sep-03	-0.0024	0.003	-0.81	0.0000	0.003	0.00	-0.0005	0.002	-0.22	-0.0031	0.003	-1.02
N	77,128			77,128			77,128			77,128		
Joint sign. test	9.685			8.277			9.720			9.900		
p-value	0.377			0.507			0.374			0.359		
Test of linear impact	9.402			8.214			5.771			4.974		
p-value	0.225			0.314			0.567			0.663		

Exhibit E2.d. Impacts on STW: Linear and IV Models with Controls for State Level Unemployment

	By Rollout Month 12			By Rollout Month 24			By Rollout Month 36			By Rollout Month 48		
	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat
LM Results for IMM: Impact on STW												
Phase 3												
Nov-03	-0.0014*	0.001	-1.95	-0.0008	0.001	-0.57	-0.0003	0.002	-0.17	-0.0000	0.001	-0.00
Jan-04	0.0024*	0.001	1.77	0.0033**	0.001	2.47	0.0051**	0.002	2.14	0.0039	0.003	1.34
Feb-04	-0.0002	0.001	-0.27	0.0012	0.001	1.18	0.0002	0.002	0.13	0.0006	0.002	0.30
Mar-04	0.0007	0.001	0.73	0.0007	0.002	0.40	-0.0006	0.002	-0.37	-0.0006	0.002	-0.29
Apr-04	-0.0001	0.001	-0.18	-0.0008	0.002	-0.43	-0.0002	0.002	-0.07	-0.0005	0.003	-0.17
May-04	-0.0004	0.001	-0.41	0.0018	0.001	1.31	0.0028	0.002	1.56	0.0054***	0.002	3.06
Jun-04	0.0020**	0.001	2.54	0.0008	0.001	0.71	0.0005	0.001	0.33	-0.0002	0.002	-0.13
Jul-04	-0.0018	0.001	-1.64	-0.0038***	0.001	-2.78	-0.0045***	0.002	-2.92	-0.0053***	0.002	-2.76
Aug-04	-0.0011	0.001	-1.09	-0.0005	0.002	-0.32	-0.0016	0.002	-0.73	-0.0007	0.003	-0.26
Sep-04	0.0001	0.001	0.05	-0.0019	0.001	-1.76	-0.0015	0.002	-0.85	-0.0027	0.001	-1.84
N	114,377			114,377			114,377			114,377		
Joint sign. test	3.328			2.542			1.977			3.236		
p-value	0.00287			0.0171			0.0616			0.00353		
Test of linear impact	3.993			2.867			1.969			3.319		
p-value	0.00150			0.0133			0.0777			0.00548		
IV Results for MM: Impact on STW												
Phase 3												
Nov-03	-0.0015*	0.001	-1.90	-0.0007	0.001	-0.50	-0.0002	0.002	-0.11	0.0001	0.002	0.07
Jan-04	0.0026*	0.001	1.79	0.0035**	0.001	2.53	0.0055**	0.003	2.18	0.0043	0.003	1.37
Feb-04	-0.0002	0.001	-0.26	0.0013	0.001	1.20	0.0003	0.002	0.15	0.0007	0.002	0.33
Mar-04	0.0008	0.001	0.74	0.0007	0.002	0.41	-0.0006	0.002	-0.36	-0.0006	0.002	-0.28
Apr-04	-0.0001	0.001	-0.19	-0.0008	0.002	-0.44	-0.0002	0.002	-0.07	-0.0005	0.003	-0.17
May-04	-0.0005	0.001	-0.42	0.0019	0.001	1.32	0.0030	0.002	1.58	0.0058***	0.002	3.09
Jun-04	0.0021**	0.001	2.53	0.0009	0.001	0.70	0.0005	0.002	0.32	-0.0002	0.002	-0.15
Jul-04	-0.0020	0.001	-1.64	-0.0041***	0.001	-2.79	-0.0048***	0.002	-2.90	-0.0058***	0.002	-2.78
Aug-04	-0.0012	0.001	-1.09	-0.0006	0.002	-0.34	-0.0018	0.002	-0.75	-0.0008	0.003	-0.28
Sep-04	0.0000	0.001	0.04	-0.0021	0.001	-1.74	-0.0017	0.002	-0.86	-0.0029	0.002	-1.84
N	114,377			114,377			114,377			114,377		
Joint sign. test	29.10			22.75			17.83			30.15		
p-value	0.000624			0.00679			0.0372			0.000414		
Test of linear impact	26.80			20.43			14.37			24.23		
p-value	0.000362			0.00472			0.0450			0.00104		

Exhibit E2.e. Impacts on NSTW Months: Linear and IV Models with Controls for State Level Unemployment

	By Rollout Month 12			By Rollout Month 24			By Rollout Month 36			By Rollout Month 48		
	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat
LM Results for IMM: Impact on NSTW Months												
Phase 1 Except NY												
Feb-02	-0.0115	0.019	-0.62	-0.0274	0.035	-0.78	-0.0352	0.053	-0.66	-0.0204	0.077	-0.26
Apr-02	0.0180	0.015	1.17	0.0433	0.029	1.49	0.0408	0.038	1.08	0.0390	0.051	0.77
May-02	-0.0036	0.010	-0.37	-0.0047	0.029	-0.16	0.0135	0.054	0.25	0.0218	0.083	0.26
Jun-02	-0.0029	0.011	-0.27	-0.0112	0.019	-0.60	-0.0191	0.035	-0.55	-0.0405	0.053	-0.77
N	43,043			43,043			43,043			43,043		
Joint sign. test	0.542			1.106			0.884			0.906		
p-value	0.656			0.356			0.456			0.445		
Test of linear impact	1.124			1.477			0.450			0.111		
p-value	0.294			0.230			0.505			0.740		
IV Results for MM: Impact on NSTW Months												
Phase 1 Except NY												
Feb-02	-0.0115	0.019	-0.62	-0.0274	0.035	-0.79	-0.0353	0.053	-0.67	-0.0204	0.076	-0.27
Apr-02	0.0181	0.015	1.18	0.0436	0.029	1.51	0.0411	0.038	1.09	0.0393	0.051	0.78
May-02	-0.0036	0.010	-0.38	-0.0048	0.029	-0.16	0.0135	0.054	0.25	0.0220	0.082	0.27
Jun-02	-0.0029	0.011	-0.27	-0.0113	0.019	-0.61	-0.0193	0.035	-0.55	-0.0409	0.053	-0.78
N	43,043			43,043			43,043			43,043		
Joint sign. test	1.661			3.391			2.709			2.776		
p-value	0.646			0.335			0.439			0.427		
Test of linear impact	1.150			1.508			0.458			0.113		
p-value	0.284			0.219			0.499			0.736		

Exhibit E2.e. Impacts on NSTW Months: Linear and IV Models with Controls for State Level Unemployment

	By Rollout Month 12			By Rollout Month 24			By Rollout Month 36			By Rollout Month 48		
	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat
LM Results for IMM: Impact on NSTW Months												
Phase 2												
Nov-02	-0.0293**	0.012	-2.39	-0.0351	0.028	-1.25	-0.0214	0.045	-0.48	-0.0199	0.069	-0.29
Jan-03	0.0031	0.010	0.31	0.0147	0.028	0.53	0.0013	0.049	0.03	-0.0102	0.071	-0.14
Feb-03	0.0031	0.008	0.38	0.0007	0.025	0.03	0.0011	0.040	0.03	0.0217	0.057	0.38
Mar-03	-0.0033	0.014	-0.24	-0.0344	0.038	-0.91	-0.0761	0.061	-1.26	-0.0938	0.088	-1.06
Apr-03	-0.0034	0.012	-0.30	-0.0230	0.026	-0.89	-0.0440	0.041	-1.07	-0.0422	0.056	-0.76
May-03	0.0037	0.012	0.31	0.0031	0.028	0.11	0.0118	0.046	0.26	0.0158	0.065	0.24
Jun-03	0.0194**	0.009	2.12	0.0332	0.023	1.42	0.0353	0.042	0.85	0.0264	0.067	0.39
Jul-03	0.0046	0.015	0.30	0.0052	0.031	0.17	0.0026	0.042	0.06	0.0053	0.051	0.10
Aug-03	0.0053	0.012	0.43	0.0155	0.031	0.50	0.0304	0.050	0.61	0.0129	0.078	0.17
Sep-03	-0.0031	0.018	-0.17	0.0201	0.042	0.48	0.0589	0.062	0.95	0.0839	0.081	1.03
N	77,128			77,128			77,128			77,128		
Joint sign. test	1.033			1.170			0.632			0.343		
p-value	0.427			0.334			0.764			0.956		
Test of linear impact	1.130			1.455			0.802			0.370		
p-value	0.359			0.204			0.590			0.916		
IV Results for MM: Impact on NSTW Months												
Phase 2												
Nov-02	-0.0306**	0.013	-2.43	-0.0370	0.029	-1.28	-0.0234	0.046	-0.51	-0.0219	0.072	-0.31
Jan-03	0.0031	0.011	0.29	0.0146	0.029	0.50	-0.0003	0.051	-0.01	-0.0125	0.074	-0.17
Feb-03	0.0031	0.009	0.36	-0.0000	0.026	-0.00	-0.0006	0.042	-0.01	0.0207	0.060	0.34
Mar-03	-0.0035	0.015	-0.24	-0.0365	0.040	-0.92	-0.0811	0.064	-1.27	-0.0998	0.093	-1.07
Apr-03	-0.0036	0.012	-0.29	-0.0246	0.027	-0.91	-0.0475	0.043	-1.10	-0.0456	0.058	-0.78
May-03	0.0040	0.013	0.31	0.0031	0.030	0.11	0.0120	0.048	0.25	0.0161	0.067	0.24
Jun-03	0.0205**	0.010	2.16	0.0351	0.024	1.45	0.0375	0.043	0.87	0.0281	0.070	0.40
Jul-03	0.0049	0.016	0.30	0.0059	0.033	0.18	0.0038	0.044	0.09	0.0068	0.055	0.13
Aug-03	0.0056	0.013	0.42	0.0171	0.034	0.51	0.0344	0.054	0.64	0.0160	0.084	0.19
Sep-03	-0.0034	0.020	-0.17	0.0222	0.046	0.49	0.0652	0.067	0.97	0.0921	0.089	1.04
N	77,128			77,128			77,128			77,128		
Joint sign. test	9.664			10.85			5.825			3.143		
p-value	0.378			0.286			0.757			0.958		
Test of linear impact	8.024			10.43			5.735			2.654		
p-value	0.330			0.165			0.571			0.915		

Exhibit E2.e. Impacts on NSTW Months: Linear and IV Models with Controls for State Level Unemployment

	By Rollout Month 12			By Rollout Month 24			By Rollout Month 36			By Rollout Month 48		
	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat	Coeff.	SE	t/z-stat
LM Results for IMM: Impact on NSTW Months												
Phase 3												
Nov-03	-0.0045	0.010	-0.48	-0.0134	0.021	-0.64	-0.0141	0.035	-0.40	-0.0248	0.048	-0.51
Jan-04	0.0156*	0.008	1.90	0.0429**	0.017	2.46	0.0834**	0.033	2.50	0.1135**	0.053	2.15
Feb-04	-0.0097	0.007	-1.41	-0.0218	0.018	-1.23	-0.0330	0.031	-1.06	-0.0295	0.049	-0.60
Mar-04	-0.0038	0.009	-0.41	-0.0006	0.021	-0.03	0.0024	0.036	0.07	-0.0231	0.047	-0.49
Apr-04	-0.0008	0.007	-0.11	-0.0100	0.020	-0.49	-0.0213	0.040	-0.54	-0.0247	0.065	-0.38
May-04	-0.0251***	0.009	-2.80	-0.0364	0.022	-1.62	-0.0345	0.035	-0.99	-0.0105	0.050	-0.21
Jun-04	0.0119	0.011	1.04	0.0223	0.027	0.84	0.0297	0.040	0.73	0.0483	0.049	0.98
Jul-04	0.0035	0.006	0.61	0.0030	0.016	0.18	-0.0137	0.024	-0.58	-0.0463	0.031	-1.48
Aug-04	0.0115	0.009	1.25	0.0120	0.021	0.56	0.0073	0.041	0.18	-0.0001	0.058	-0.00
Sep-04	0.0013	0.007	0.20	0.0021	0.016	0.13	-0.0063	0.026	-0.24	-0.0026	0.038	-0.07
N	114,377			114,377			114,377			114,377		
Joint sign. test	5.035			2.137			2.144			2.446		
p-value	7.31e-05			0.0430			0.0423			0.0213		
Test of linear impact	5.007			2.620			2.172			2.456		
p-value	0.000228			0.0217			0.0524			0.0299		
IV Results for MM: Impact on NSTW Months												
Phase 3												
Nov-03	-0.0050	0.010	-0.49	-0.0143	0.022	-0.64	-0.0145	0.037	-0.39	-0.0255	0.051	-0.50
Jan-04	0.0164*	0.009	1.90	0.0453**	0.018	2.47	0.0884**	0.035	2.51	0.1204**	0.056	2.16
Feb-04	-0.0104	0.007	-1.42	-0.0232	0.019	-1.24	-0.0349	0.033	-1.06	-0.0311	0.051	-0.60
Mar-04	-0.0041	0.010	-0.42	-0.0007	0.022	-0.03	0.0026	0.038	0.07	-0.0245	0.050	-0.49
Apr-04	-0.0009	0.007	-0.12	-0.0107	0.021	-0.50	-0.0228	0.042	-0.54	-0.0264	0.069	-0.38
May-04	-0.0269***	0.010	-2.82	-0.0390	0.024	-1.64	-0.0369	0.037	-1.00	-0.0112	0.053	-0.21
Jun-04	0.0128	0.012	1.05	0.0239	0.028	0.84	0.0318	0.043	0.74	0.0516	0.052	0.99
Jul-04	0.0039	0.006	0.63	0.0033	0.018	0.19	-0.0147	0.025	-0.58	-0.0500	0.033	-1.50
Aug-04	0.0125	0.010	1.27	0.0130	0.023	0.57	0.0078	0.044	0.18	-0.0003	0.062	-0.01
Sep-04	0.0016	0.007	0.22	0.0024	0.018	0.14	-0.0068	0.028	-0.24	-0.0030	0.041	-0.07
N	114,377			114,377			114,377			114,377		
Joint sign. test	48.28			20.05			20.08			23.40		
p-value	2.26e-07			0.0176			0.0175			0.00536		
Test of linear impact	37.06			19.15			16.00			18.37		
p-value	4.56e-06			0.00772			0.0251			0.0104		

Note:
Tables E2a-e do not include results for Phase 1 NY sample. The variables based on state level unemployment rates get dropped from the analysis due to collinearity. Consequently, the model specification for Phase 1 NY sample becomes the same as the specification without the state level unemployment variables, results for which are presented in Tables E1a-e.

This page has been left blank for double-sided copying.

APPENDIX F

DEFINITIONS OF VARIABLES AND SUMMARY DESCRIPTION OF MODELS

This page has been left blank for double-sided copying.

Exhibit F.1. List of Variables Used in the Analysis[49]

Variable Name	Variable Label/Description
Outcome Variables	
srvroll12, srvroll24, srvroll36, srvroll48	Enrolled in services within 12/24/36/48 months following start of rollout
twproll12, twproll24, twproll36, twproll48	TWP began within 12/24/36/48 months following start of rollout
eperoll12, eperoll24, eperoll36, eperoll48	TWP completed within 12/24/36/48 months following start of rollout
ldwroll12, ldwroll24, ldwroll36, ldwroll48	Suspended/terminated for work (aka, left due to work) within 12/24/36/48 months following start of rollout
nstw12, nstw24 nstw36, nstw48	Number of months suspended/terminated for work within 12/24/36/48 months following start of rollout
Intended Ticket Mail Indicators	
Frstmldt	first ticket mailing date
imm1	=1 if terminal digit of SSN is 1 and a phase 1 beneficiary
imm2	no tickets were intended to be mailed in month 2
imm3	=1 if terminal digit of SSN is 2 or 3 and a phase 1 beneficiary not from NY
imm4	=1 if terminal digit of SSN is 4, 5 or 6 and a phase 1 beneficiary not from NY or if terminal digit of SSN is 2 and phase 1 beneficiary from NY
imm5	=1 if terminal digit of SSN is 7, 8, 9 or 0 and a phase 1 beneficiary not from NY
imm6	=1 if terminal digit of SSN is 3 or 4 and a phase 1 beneficiary from NY
imm7	=1 if terminal digit of SSN is 5 or 6 and a phase 1 beneficiary from NY
imm8	=1 if terminal digit of SSN is 7 or 8 and a phase 1 beneficiary from NY
imm9	=1 if terminal digit of SSN is 9 or 0 and a phase 1 beneficiary from NY
imm10	=1 if terminal digit of SSN is 1 and a phase 2 beneficiary
imm11	no tickets were intended to be mailed in month 11
imm12	=1 if terminal digit of SSN is 2 and a phase 2 beneficiary
imm13	=1 if terminal digit of SSN is 3 and a phase 2 beneficiary
imm14	=1 if terminal digit of SSN is 4 and a phase 2 beneficiary
imm15	=1 if terminal digit of SSN is 5 and a phase 2 beneficiary
imm16	=1 if terminal digit of SSN is 6 and a phase 2 beneficiary
imm17	=1 if terminal digit of SSN is 7 and a phase 2 beneficiary
imm18	=1 if terminal digit of SSN is 8 and a phase 2 beneficiary
imm19	=1 if terminal digit of SSN is 9 and a phase 2 beneficiary
imm20	=1 if terminal digit of SSN is 0 and a phase 2 beneficiary
imm21	=1 if terminal digit of SSN is 1 and a phase 2 beneficiary
imm22	no tickets were intended to be mailed in month 22
imm23	=1 if terminal digit of SSN is 2 and a phase 3 beneficiary
imm24	=1 if terminal digit of SSN is 3 and a phase 3 beneficiary
imm25	=1 if terminal digit of SSN is 4 and a phase 3 beneficiary
imm26	=1 if terminal digit of SSN is 5 and a phase 3 beneficiary
imm27	=1 if terminal digit of SSN is 6 and a phase 3 beneficiary
imm28	=1 if terminal digit of SSN is 7 and a phase 3 beneficiary
imm29	=1 if terminal digit of SSN is 8 and a phase 3 beneficiary
imm30	=1 if terminal digit of SSN is 9 and a phase 3 beneficiary
imm31	=1 if terminal digit of SSN is 0 and a phase 3 beneficiary
motoimm	Number of months between ticket selection date and intended mail month.
mototkt	Number of months between rollout date and ticket mail date.

[49] The acronyms used for TWP completion and STW was different in the codes, variable names, and file names included in Appendix F and Appendix G: the shorthand used for TWP completion is "EPE" (for extended period of eligibility, which coincides with the TWP completion plus a three month grace period); the shorthand used for STW is "LDW" (for left due to work).

Table F.1 *(continued)*

Variable Name	Variable Label/Description
Control and Other Variables	
awarddate	DI award date
phase	Phase of Ticket to Work program rollout
male	Male
gendermiss_flag	Variable to identify records with gender missing
awardage2	Age at DI award
doage2	Age at disability onset
doage2miss_flag	Variable to identify records with age at disability onset missing
tsd_age	Age at ticket selection date
race_a	Asian
race_b	African American
race_h	Hispanic
race_i	Native American
race_w	White
race_o	Other race
race_mis	Missing race
tsd_edu_hs	12 yrs of education at ticket selection date
tsd_edu_mrhs	>12 yrs of education at ticket selection date
tsd_edu_mis	Missing education info
tsd_mie_psbl	Medical improvement possible at ticket selection date
tsd_mie_exp	Medical improvement expected at ticket selection date
tsd_mie_mis	Missing medical improvement info at ticket selection date
tsd_mie_ne	Medical improvement not expected at ticket selection date
tsd_medicare	Eligible for Medicare at ticket selection date
tsd_medicare_miss	Missing Medicare info at ticket selection date
tsd_depend_nil	No dependents at ticket selection date
tsd_depend_1	1 dependents at ticket selection date
tsd_depend_2	>2 dependents at ticket selection date
tsd_depend_miss	Missing dependent info at ticket selection date
tsd_vrpr	Eligible for VR services prior ticket selection date
tsd_vrpr_miss	Missing information about VR eligibility prior to ticket selection date
pdcgroup1	Primary disabling condition: major affective disorders
pdcgroup2	Primary disabling condition: other psychiatric disorders and mental retardation
pdcgroup3	Primary disabling condition: back disorders and musculoskeletal system
pdcgroup4	Primary disabling condition: other physical disabilities
pdcgroup5	Primary disabling condition: missing
cohort1999 – cohort2005	Year of DI award 1999/2000/…/2005
tsd_unemp_mean	Average state unemployment rate from two months before ticket selection date to three months after ticket selection date
tsd_unemp_cng	Change in state unemployment rate from two months before ticket selection date to three months after ticket selection date
pia1	First instance of primary insurance amount from the DAF
pia_miss	Missing PIA1
ime1	First instance of indexed monthly earnings from the DAF
ime_miss	Missing IME1
award_b4_tsd	DI award date before ticket selection date
diaward_tsd	Number of months between DI award date and ticket selection date
epeb4twp_flag	TWP completion date is before TWP start date
ldwb4twp_flag	suspended/terminated for work (aka, left due to work) before beginning TWP
ldwb4epe_flag	Suspended/terminated for work (aka, left due to work) before TWP completion
twpb4tsd	TWP began before ticket selection date
epeb4tsd	TWP completed before ticket selection date
ldwb4tsd	Suspended/terminated for work (aka, left due to work) before ticket selection date
st_AK	state == AK
st_AL	state == AL
st_AR	state == AR
st_AZ	state == AZ
st_CA	state == CA
st_CO	state == CO
st_CT	state == CT
st_DC	state == DC

Table F.1 *(continued)*

Variable Name	Variable Label/Description
st_DE	state == DE
st_FL	state == FL
st_GA	state == GA
st_HI	state == HI
st_IA	state == IA
st_ID	state == ID
st_IL	state == IL
st_IN	state == IN
st_KS	state == KS
st_KY	state == KY
st_LA	state == LA
st_MA	state == MA
st_MD	state == MD
st_ME	state == ME
st_MI	state == MI
st_MN	state == MN
st_MO	state == MO
st_MS	state == MS
st_MT	state == MT
st_NC	state == NC
st_ND	state == ND
st_NE	state == NE
st_NH	state == NH
st_NJ	state == NJ
st_NM	state == NM
st_NV	state == NV
st_NY	state == NY
st_OH	state == OH
st_OK	state == OK
st_OR	state == OR
st_PA	state == PA
st_PR	state == PR
st_RI	state == RI
st_SC	state == SC
st_SD	state == SD
st_TN	state == TN
st_TR	state == Territories
st_TX	state == TX
st_UT	state == UT
st_VA	state == VA
st_VT	state == VT
st_WA	state == WA
st_WI	state == WI
st_WV	state == WV
st_WY	state == WY
st_ZZ	state == missing

Exhibit F.2. Summary of Analytic Models

Analysis Using Random Within-State Variation in Mail Month to Identify Impacts	
1. Linear Probability Model—Discrete Intended Mail Months (IMM)	
Software	Stata 11.2
Model description	Linear probability models for the probability that the event has occurred at 12, 24, 36 and 48 months following rollout start by phase, with discrete indicators for the IMM
Procedure	reg `e'roll`n' (intended mail month indicators) male gendermiss_flag tsd_age doage2 doage2miss_flag race_a race_b race_h race_i race_o race_mis tsd_edu_hs tsd_edu_mrhs tsd_edu_mis tsd_mie_exp /*tsd_mie_ne */ tsd_mie_mis tsd_mie_psbl tsd_medicare tsd_medicare_miss tsd_depend_1 tsd_depend_2 tsd_depend_miss tsd_vrpr tsd_vrpr_miss pdcgroup2 pdcgroup3 pdcgroup4 pdcgroup5 /*cohort1999*/ cohort2000 cohort2001 cohort2002 cohort2003 cohort2004 award_b4_tsd diaward_tsd epeb4twp_flag ldwb4twp_flag ldwb4epe_flag twpb4tsd epeb4tsd ldwb4tsd st_AL-st_TN st_TX-st_WY tsd_unemp_mean tsd_unemp_cng pia1 pia_miss ime1 ime_miss, vce(cluster tsd_state) Where `e'roll`n' = srvroll`n', twproll`n', eperoll`n', ldwroll`n', nstw`n' and `n' = 12, 24, 36, 48.
Options used	Used the "vce (cluster *clustervar*)" option to correct estimated standard errors for heteroskedasticity and clustering at the state level; for Phase 1 NY sample, used the "robust" option so that the estimated standard errors are heteroskedasticity robust
Variation in model specification	We ran two versions of the model specification: one that included measures of state level unemployment rates (tsd_unemp_mean tsd_unemp_cng), and another that did not.
Program and output files	2. Analysis\A. Linear probability models\ Program file: LPM_ModelA.do Log file: LPM_ModelA.txt Outreg output: \LPMOutput\1. Discrete Intended Mail Months\ LPM_PHx_unemp.xls (or .txt) LPM_PHx_nounemp.xls (or .txt) where x = 1NONY, 1NY, 2, and 3
2. Linear Probability Model—Continuous Intended Mail Months (IMM)	
Software	Stata 11.2
Model description	Linear probability models for the probability that the event has occurred at 12, 24, 36 and 48 months following rollout start by phase, with continuous measure of duration between rollout start and the IMM
Procedure	reg `e'roll`n' (duration to intended mail month) male gendermiss_flag tsd_age doage2 doage2miss_flag race_a race_b race_h race_i race_o race_mis tsd_edu_hs tsd_edu_mrhs tsd_edu_mis tsd_mie_exp /*tsd_mie_ne */ tsd_mie_mis tsd_mie_psbl tsd_medicare tsd_medicare_miss tsd_depend_1 tsd_depend_2 tsd_depend_miss tsd_vrpr tsd_vrpr_miss pdcgroup2 pdcgroup3 pdcgroup4 pdcgroup5 /*cohort1999*/ cohort2000 cohort2001 cohort2002 cohort2003 cohort2004 award_b4_tsd diaward_tsd epeb4twp_flag ldwb4twp_flag ldwb4epe_flag twpb4tsd epeb4tsd ldwb4tsd st_AL-st_TN st_TX-st_WY tsd_unemp_mean tsd_unemp_cng pia1 pia_miss ime1 ime_miss, vce(cluster tsd_state) Where `e'roll`n' = srvroll`n', twproll`n', eperoll`n', ldwroll`n', nstw`n' and `n' = 12, 24, 36, 48.
Options used	Used the "vce (cluster *clustervar*)" option to correct estimated standard errors for heteroskedasticity and clustering at the state level; for Phase 1 NY sample, used the "robust" option so that the estimated standard errors are heteroskedasticity robust
Variation in model specification	We ran two versions of the model specification: one that included measures of state level unemployment rates (tsd_unemp_mean tsd_unemp_cng), and another that did not.
Program and output files	2. Analysis\A. Linear probability models\ Program file: LPM_ModelC.do Log file: LPM_ModelCtxt Outreg output:

Table F.2 *(continued)*

Analysis Using Random Within-State Variation in Mail Month to Identify Impacts	
	\LPMOutput\2. Continuous Intended Mail Months\ LPM_PHx_unemp.xls (or .txt) LPM_PHx_nounemp.xls (or .txt) where x = 1NONY, 1NY, 2 and 3
3. Instrumental Variables Model—Discrete Mail Months (MM)	
Software	Stata 11.2
Model description	Instrumental variables models for the probability that the event has occurred at 12, 24, 36 and 48 months following rollout start by phase, using the discrete IMM indicators as instruments for the discrete indicators of the actual mail months (MM)
Procedure	ivreg2 `e'roll`n' (ph`x'mm_adj'= `ph`x'imm_adj) male gendermiss_flag tsd_age doage2 doage2miss_flag race_a race_b race_h race_i race_o race_mis tsd_edu_hs tsd_edu_mrhs tsd_edu_mis tsd_mie_exp /*tsd_mie_ne */ tsd_mie_mis tsd_mie_psbl tsd_medicare tsd_medicare_miss tsd_depend_1 tsd_depend_2 tsd_depend_miss tsd_vrpr tsd_vrpr_miss pdcgroup2 pdcgroup3 pdcgroup4 pdcgroup5 /*cohort1999*/ cohort2000 cohort2001 cohort2002 cohort2003 cohort2004 award_b4_tsd diaward_tsd epeb4twp_flag ldwb4twp_flag ldwb4epe_flag twpb4tsd epeb4tsd ldwb4tsd st_AL-st_TN st_TX-st_WY tsd_unemp_mean tsd_unemp_cng pia1 pia_miss ime1 ime_miss, ffirst partiaul (`covar') cluster (tsd_state) Where `e'roll`n' = srvroll`n', twproll`n', eperoll`n', ldwroll`n', ,nstw`n' and `n' = 12, 24, 36, 48.
Options used	Used the "cluster (*clustervar*)" option to correct estimated standard errors for heteroskedasticity and clustering at the state level; for Phase 1 NY sample, used the "robust" option so that the estimated standard errors are heteroskedasticity robust Used the "ffirst" option to report only these identification statistics and not the first-stage regression results themselves Used the "partial" option to partial out some of the regressors from the the first stage equation
Variation in model specification	We ran two versions of the model specification: one that included measures based on state level unemployment rates (tsd_unemp_mean tsd_unemp_cng), and another that did not.
Program and output files	2. Analysis\B. Instrumental variables models\ Program file: IV_ModelD.do Log file: IV_ModelD.txt Outreg output: \LPMOutput\1. Discrete Mail Months\ LPM_PHx_unemp.xls (or .txt) LPM_PHx_nounemp.xls (or .txt) where x = 1NONY, 1NY, 2 and 3
4. Instrumental Variables Model—Continuous Mail Months (MM)	
Software	Stata 11.2
Model description	Instrumental variables models for the probability that the event has occurred at 12, 24, 36 and 48 months following rollout start by phase, using the discrete IMM indicators as instruments for the duration between rollout start and the actual mail month (MM)
Procedure	ivreg2 `e'roll`n' (mototkt'= `ph`x'imm_adj) male gendermiss_flag tsd_age doage2 doage2miss_flag race_a race_b race_h race_i race_o race_mis tsd_edu_hs tsd_edu_mrhs tsd_edu_mis tsd_mie_exp /*tsd_mie_ne */ tsd_mie_mis tsd_mie_psbl tsd_medicare tsd_medicare_miss tsd_depend_1 tsd_depend_2 tsd_depend_miss tsd_vrpr tsd_vrpr_miss pdcgroup2 pdcgroup3 pdcgroup4 pdcgroup5 /*cohort1999*/ cohort2000 cohort2001 cohort2002 cohort2003 cohort2004 award_b4_tsd diaward_tsd epeb4twp_flag ldwb4twp_flag ldwb4epe_flag twpb4tsd epeb4tsd ldwb4tsd st_AL-st_TN st_TX-st_WY tsd_unemp_mean tsd_unemp_cng pia1 pia_miss ime1 ime_miss, ffirst partiaul (`covar') cluster (tsd_state) Where `e'roll`n' = srvroll`n', twproll`n', eperoll`n', ldwroll`n', and `n' = 12, 24, 36, 48.
Options used	Used the "cluster (*clustervar*)" option to correct estimated standard errors for heteroskedasticity and clustering at the state level; for Phase 1 NY sample, used the "robust" option so that the estimated standard errors are heteroskedasticity

Table F.2 *(continued)*

Analysis Using Random Within-State Variation in Mail Month to Identify Impacts	
	robust Used the "ffirst" option to report only these identification statistics and not the first-stage regression results themselves Used the "partial" option to partial out some of the regressors from the the first stage equation
Variation in model specification	We ran two versions of the model specification: one that included measures based on state level unemployment rates (tsd_unemp_mean tsd_unemp_cng), and another that did not.
Program and output files	2. Analysis\B. Instrumental variables models\ Program file: IV_ModelC.do Log file: IV_ModelC.txt Outreg output: \LPMOutput\2. Continuous Mail Months\ LPM_PHx_unemp.xls (or .txt) LPM_PHx_nounemp.xls (or .txt) where x = 1NONY, 1NY, 2 and 3
5. Projections for Total Impacts	
Software	Stata 11.2
Model description	Used instrumental variables estimates to project total impacts at 12, 24, 36 and 48 month after ticket mailing.
Procedure	reg3 (`v'12 mototkt `nounempny') /// (`v'24 mototkt `nounempny') /// (`v'36 mototkt `nounempny') /// (`v'48 mototkt `nounempny') if phase1_st_ny, endog(mototkt) exog(`ph1nyimm') Where `v' = srvroll, twproll, eperoll, ldwroll, and nstw. Where `nounempny' = male gendermiss_flag tsd_age doage2 doage2miss_flag race_a race_b race_h race_i race_o race_mis tsd_edu_hs tsd_edu_mrhs tsd_edu_mis tsd_mie_exp /*tsd_mie_ne */ tsd_mie_mis tsd_mie_psbl tsd_medicare tsd_medicare_miss tsd_depend_1 tsd_depend_2 tsd_depend_miss tsd_vrpr tsd_vrpr_miss pdcgroup2 pdcgroup3 pdcgroup4 pdcgroup5 /*cohort1999*/ cohort2000 cohort2001 cohort2002 cohort2003 cohort2004 award_b4_tsd diaward_tsd epeb4twp_flag ldwb4twp_flag ldwb4epe_flag twpb4tsd epeb4tsd ldwb4tsd st_AL-st_TN st_TX-st_WY tsd_unemp_mean tsd_unemp_cng pia1 pia_miss ime1 ime_miss
Program and output files	2. Analysis\C. Projections for total impacts\ Program file: 3sls.do Log file: 3sls.txt Outreg output: 3slsOutput\3sls_output.xlsx

This page has been left blank for double-sided copying.

MATHEMATICA
Policy Research

www.mathematica-mpr.com

Improving public well-being by conducting high quality, objective research and surveys

Princeton, NJ ■ Ann Arbor, MI ■ Cambridge, MA ■ Chicago, IL ■ Oakland, CA ■ Washington, DC