Disease Management: Helping Patients (Who Don't) Help Themselves *

Paul Gertler
University of California at Berkeley, Haas School of Business and NBER

Tim Simcoe
University of Toronto, Joseph L. Rotman School of Management

April 2009

*Thanks to Derek Newell, Yang Gao, Alex Exuzides and LifeMasters Supported Self Care, Inc. for providing data and insights into LifeMasters' disease management program. We also received useful comments from Meghan Busse, Ken Chay, Catherine Wolfram, Florian Zettlemeyer and participants in Spring 2004 NBER health and productivity meetings. Address for correspondence: J.L. Rotman School of Management, University of Toronto, Toronto, ON, M5S 3E6. E-mail: timothy.simcoe@rotman.utoronto.ca

Disease Management: Helping Patients (Who Don't) Help Themselves

Abstract

Chronically ill patients currently consume a significant share of the U.S. health system's resources and are a rapidly growing segment of the overall population. Disease Management (DM) programs identify high-risk patients among the chronically ill, encourage them to take better care of themselves, and help coordinate the care they receive from various providers. This paper examines the impact of a diabetes Disease Management program. We find that it led to increased compliance with clinical practice guidelines, improvements in patient health, and significant reductions in the total cost of care. The financial benefits are greater for patients lacking "self control" prior to enrollment, as indicated by their failure to comply with generally accepted clinical practice guidelines. These results are especially important for the Medicare program, which has the majority of the chronically ill as beneficiaries.
Keywords: Disease management, self control, diabetes. *JEL Codes*: I12.

1 Introduction

Patients with chronic illnesses, such as coronary artery disease, congestive heart failure, and diabetes, consume almost three-quarters of the $1.6 trillion spent annually on medical care in the United States (Hoffman and Rice 1995; Levit, Smith, Cowan et al 2004). These costs are projected to grow as the population ages and as the obesity epidemic expands (Sturm, Ringel and Andreyeva 2004), and Medicare will bear much of the burden (Medicare Payment Advisory Commission 2005).

The size and scope of this problem has led to a broad search for methods of improving health outcomes that also reduce the costs of medical care for the chronically ill. In 2001, The Institute of Medicine (IOM) of the US National Academies of Science identified Disease Management (DM) as a promising but untested solution (IOM 2001). Disease management programs are a form of preventive care designed to reduce the likelihood that a chronic illness leads to costly complications. DM programs encourage patients to live a healthy lifestyle and follow their prescribed treatments through monitoring and regular contact. In particular, they facilitate patient compliance with IOM clinical practice guidelines (CPGs), which are consensus evidence-based disease-specific preventive and treatment activities that maximize the probability of staying healthy.

While this sounds straightforward, over half of chronically ill patients do not achieve adequate compliance levels (NCQA 2005). Failure to adhere to these simple guidelines frequently leads to a loss of control over the medical condition, resulting in serious complications that require expensive hospitalizations and procedures. The goal of DM is to prevent these adverse outcomes by identifying high-risk patients and encouraging them to take an active role in managing their own health. Disease Management encourages CPG compliance in two ways: first by educating patients about the benefits of compliance and the costs of non-compliance, and second by supplying one-on-one help to those that do not have enough "self control" to help themselves.

Can DM programs actually produce enough change in patient behaviors to significantly improve health outcomes and reduce medical costs? To answer this question, we examine the impact of a large disease management program for diabetes. Diabetes is the most common chronic disease. In fact, 4.2 percent of the population has diabetes and the prevalence is expected to grow to 5.2 percent in the next 15 years (Hogan, Dall and Nikolov 2003). Medical care expenditures by patients with diabetes are roughly $160 billion per year. And the growth in diabetes has critical implications for Medicare, as 52 percent of the population with diabetes are over age 65.

We take advantage of longitudinal data on a population of beneficiaries to obtain difference

in difference estimates of the impacts of the DM program. Our results show that it led to increased CPG compliance and produced significant health improvements—measured in terms of blood sugar levels, hospitalization and the onset of serious complications. We also find that it produced a statistically and economically significant decline in medical expenditures. However, all of the financial benefits of the program were concentrated among patients who exhibited low "self control" before enrolling in the DM program (i.e. they did not comply with clinical practice guidelines).

Our estimates indicate that this DM program generated roughly $1 million in savings during its first three years—more than 5 times the cost of the program. If these results hold for the broader population of persons with diabetes, rough calculations suggest that DM has the potential to reduce the annual cost of care for diabetes in the US by $7.6 billion.[1] Medicare alone could save roughly $2 billion annually.[2]

The remainder of the paper proceeds as follows. The next section discusses the economics of disease management and preventive care. Section 3 describes the diabetes DM program that we evaluate below. Section 4 outlines our methods, and Section 5 describes the data and summary statistics. Section 6 presents empirical results, and Section 7 concludes.

2 The Economics of Disease Management and Self-Control

Disease management is a set of practices designed to encourage and assist chronically ill patients' efforts to look after their own health. DM programs target the most common chronic illnesses, such as asthma, diabetes, and congestive heart failure. They use a combination of patient and physician education, personal coaching by nurse case-managers, and information technology-based monitoring of patient compliance and outcomes.

Disease management exists because not everyone takes care of their health. This observation applies to a wide range of behaviors, from failing to exercise or maintain a healthy diet to actively engaging in harmful activities like smoking, drug use, or excessive alcohol consumption. For the chronically ill, taking care of oneself also implies a treatment regime that includes daily medications, regular diagnostic tests and visits to the doctor (as well as maintaining a healthy lifestyle). A substantial body of evidence links unhealthy behaviors and the lack of preventive

[1] Fourteen percent of all persons with diabetes in our study were eligible to enroll (i.e. were "high risk" patients) and also had poor self-control. Making the conservative assumption that these patients represent 14 percent of the $160 billion in direct spending for diabetes, and applying an estimated cost savings of 34 percent (the lower end of the 95 percent confidence interval for our main result in Table 7), we arrive at $7.62 billion.

[2] Take the previous figure and multiply by the average Medicare reimbursement rate (51.8 percent; Medicare Payment Advisory Commission, 2005), and the percentage of persons with diabetes over 65 years of age (52 percent; Hogan, Dall and Nikolov 2003).

care to mortality rates and health care costs (e.g. NCQA 2005; McGinness and Foege 1993; Chandra, Gruber and McKnight 2007; McWilliams et al 2007).

The standard economic model of preventive care views healthy behavior as an investment (Grossman 1972). In this framework differences in behavior are explained by variation in individual tastes or discount factors. Cutler and Glaeser (2005) note that one implication of this model is that "investments" in healthy behavior should be highly correlated within individuals, since they are driven by a common set of underlying parameters. Using survey data, they show that this prediction is not true. In data from several large surveys, individuals exhibit very little correlation in either their choice of unhealthy behaviors (e.g. smoking, drinking, and exercise) or changes in these behaviors.

A leading alternative to the standard "investment model" focuses on the importance of information. The basic idea is that heterogeneity in health-related behaviors may be explained by idiosyncratic differences in patient knowledge. For example, Goldman and Smith (2002) find that HIV and diabetes patients with more education—who are presumably better at acquiring and processing information—were better at managing their own care. Similarly, Rothman et al (2004) find that patients with low literacy levels benefit more from a diabetes DM intervention.

To the extent that the relevant knowledge is disease- or behavior-specific, informational theories will predict lower correlations in (un)healthy behavior. However, they cannot explain why some behaviors, such as smoking or poor diet and exercise, are remarkably persistent even though, as survey evidence suggests, their long-term health effects are widely understood. Behavioral theories provide a second alternative to viewing healthy behavior as an investment, and generally focus on explaining individuals' lack of "self-control" (i.e. the remarkable persistence of many destructive behaviors). These models either allow present utility to vary with past consumption (Becker and Murphy 1988), or assume that discounting is not exponential (Laibson 1997, O'Donohughe and Rabin 1999).

In behavioral models, idiosyncratic shocks that lead to experimentation with "bad" behaviors may have long-run consequences—thus explaining the lack of a strong cross-sectional correlation in (un)healthy behavior. Behavioral theories also predict that the relatively small cost of complying with self-care guidelines might cause otherwise well-informed and conscientious individuals to neglect their health for long periods of time.

Disease management programs are designed to address both informational and behavioral antecedents of unhealthy behavior. In particular, they address informational problems by placing a heavy emphasis on education—particularly at the start of the program. Patient education is especially important when medical knowledge is advancing rapidly, leading to complex disease-specific treatment regimes with many dimensions (e.g. dietary restrictions, medication, and self-monitoring). Over the long term, DM works to increase the psychological

cost of unhealthy behavior by continually reminding enrolled patients to look after themselves and encouraging them to set specific targets. Regular contact with patients helps those with less self-discipline to make simple lifestyle changes that are widely understood to have a significant health impact.

Most DM programs begin with a review of the health plan's medical claims data to identify a group of "high-risk" chronically ill patients. This search focuses on patients who have been hospitalized, developed serious complications, or failed to comply with standard treatment protocols. With the permission of a primary care physician, these patients are asked to enroll in the DM program free of charge. Those who accept begin by participating in an educational session focused on their disease and the importance of both life-style and therapy compliance. This is followed by a series of regularly scheduled contacts—either by phone or in person—with nurse-coaches who monitor the patient's health and provide reminders about compliance. Over time, the DM program uses claims data and IT systems (e.g. automated call centers) to monitor and interact with enrolled patients. In more sophisticated DM programs, comprehensive patient-tracking systems are used to enhance the coordination among primary care physicians, specialists, pharmacists, and the patient.

Evidence-based Clinical Practice Guidelines (CPGs) are a key input in this process. CPGs are a set of protocols for treating common diseases that DM programs use as a reference point to identify patient- or physician-specific deviations from a recommended course of treatment.[3] The existence of an established set of "best practices" enhances the credibility of DM providers, eliminating the need to develop proprietary benchmarks or to convince clients, doctors, and patients that they are clinically appropriate. These benchmarks are used to analyze claims data to assess patient CPG compliance, and to provide feedback to physicians and patients for the purposes of behavioral change and early intervention.

Disease management programs are typically paid for by insurers or HMOs, rather than physician groups or individual employers. Health plans' existing claim data bases, IT infrastructure, and scale place them in the best position to gather all of the information required to coordinate a patient's care (e.g. physician visits, lab results, and pharmacy records). Moreover, physicians and hospitals have little financial incentive to invest in preventing chronic disease complications since they are usually paid on a fee-for-service basis. Health plans generally receive a fixed per-member payment each month. Thus, as long as patient turnover rates are not too high, the plans will capture a share of any long-run cost reductions.

Health plans generally administer DM programs in one of three ways. In the "integrated

[3] Clinical practice guidelines were developed in the 1990s in response to research documenting large geographic variations in clinical practice that could not be explained by differences in epidemiology or local demographics (Wennberg and Gittelsohn 1973).

delivery system" model, a health plan develops and manages all aspects the program internally. In the "carve out" model, a third-party DM provider administers the program and takes on the financial risk associated with enrolled patients. Finally, in the "carve in" model, a third-party DM vendor works with a health plan to coordinate care for chronically ill patients, but does not assume any of the financial risks.

The direct costs of DM include an initial investment in technology and expertise, along with the ongoing cost of staffing a program. In general, the IT systems used to screen for eligibility, monitor compliance, and issue reminders are not particularly novel or sophisticated—especially for larger HMOs. The ongoing costs are primarily staff, which can include analysts, case workers, nurse-coaches, and call-center personnel. Many of these costs vary with the scale and scope of the intervention. In their 2004 Annual Report, American Healthways—the largest publicly owned third-party DM provider—indicates direct costs of $117.20 per person per year to manage 1,335,000 chronically ill patients. Beaulieu, Cutler and Ho (2002) estimate about $45 per patient-year based on a case study of two HMOs' diabetes DM programs.

While there is a substantial literature on Disease Management, the evidence on its financial impact is mixed. Early studies (e.g. those reviewed in Knight 2005) generally showed modest improvements in CPG compliance and clinical outcomes, such as HbA1c levels. However, a literature review by the Congressional Budget Office (CBO 2004) concludes that "much of the literature on [DM] programs does not directly address health care costs." Recent studies, such as Berg and Wadhwa (2007) and Bray et al (2008), have found larger clinical effects across a broad range of outcomes, including financially relevant measures such as inpatient bed-days. However, using methods similar to ours, Lairson et al (2008) find small and statistically insignificant financial impacts over a one-year post-enrollment period.

In spite of this mixed evidence, DM practices have been gaining in popularity among both public and private insurers.[4] Felt-Lisk and Mays (2002) report that almost all of the 48 private health-plans that they studied had adopted or expanded some type of DM program during the late 1990s. According to Gillespie and Rossiter (2003), more than 20 state-run Medicaid programs had adopted some type of DM program by 2002. At the federal level, the Balanced Budget Act of 1997 expanded Medicare coverage for diabetes self-education. And in 2003, Congress authorized a $100 million DM pilot program as part of the Medicare Prescription Drug Act.

[4]In contrast, Casalino et al(2003) report little adoption of a variety of chronic care management practices by independent physician groups.

3 The Intervention

We examine a diabetes DM carve-in intervention for beneficiaries of Fallon Community Health Plan, a large central Massachusetts HMO that provides health insurance for more than 214,000 members. The DM program was managed by LifeMasters, a private company specializing in chronic care management. This section describes the treatment of diabetes and provides an overview of the intervention.

Diabetes is a disease that limits the body's ability to produce and/or properly utilize insulin. Diabetes comes in two forms. Type 1 diabetes—in which the body produces no insulin—occurs primarily in children and young adults, accounting for between 5 and 10 percent of all cases. The remaining 90 to 95 percent of patients have Type 2 diabetes. All persons with diabetes have difficulty breaking down sugars and starches and must pay close attention to the level of sugars in their body. While its causes are not well understood, both genetic and environmental factors—particularly obesity and a lack of exercise—can play a role in the onset of Type 2 diabetes. Patients with diabetes have a significantly greater chance of developing a number of medical complications including blindness, kidney disease, nerve damage, heart-disease, and stroke.

As with any chronic disease, there is no cure for diabetes. Treatment is generally coordinated by a patient's primary care physician, who prescribes a regimen of diet, exercise and regular diagnostic tests designed to prevent the onset of serious complications. In most cases, the patient also has a daily regimen of prescription drugs, such as insulin or other oral medications. Between visits, diabetic patients monitor their blood glucose (sugar) levels on a regular basis and work to keep them within recommended ranges.

The most important measure of a patient's control over diabetes is based on a blood test for Hemoglobin A1C (HbA1c). This test indicates how well a patient has controlled their blood sugar levels over a three month period. An HbA1c score less than 7.0 is considered good, while a score greater than 9.5 suggests that a patient has poor control over their condition.[5]

The National Committee for Quality Assurance (NCQA) Diabetes Quality Improvement Project has developed a set of clinical guidelines for treatment of diabetes based on a large body of clinical research, including two large trials—The Diabetes Control and Complications Trial and the UK Prospective Diabetes Study Group. The following guidelines are used by organizations such as NCQA to assess the overall quality of care received by the population with diabetes. They are also the standard measures of compliance and health used by most DM programs.

[5]There is some recent (and controversial) evidence from a large clinical trial that intensive treatment designed to bring HbA1c levels below 6.0 may increase mortality among patients with Type 2 diabetes (ACCORD 2008).

1. Annual Hemoglobin A1c (HbA1c) test
2. Annual Retinal (eye) examination
3. Annual Lipids (cholesterol) profile
4. Annual monitoring for nephropathy (kidney disease)
5. Poor HbA1c control (HbA1c$\geq$9.5%)
6. Lipids controlled (LDL-C<130 mg/dL)

Surveys administered by the NCQA (2005) show that while most health plans have reasonably high compliance rates on screening for HbA1c and cholesterol (i.e. above 70 percent), compliance rates for the other four guidelines are all below 50 percent.

In May 1999, Fallon Community Health Plan contracted with LifeMasters to introduce a disease management program for patients with diabetes. At the start of the intervention, LifeMasters used medical claims data to identify patients with diabetes and group them into three risk categories. Patients were eligible to enroll in the program if they met one of the following criteria, which placed them in the high-risk category: hospitalization for a diabetes-related complication, an HbA1c level above 9.5, or an established diabetic complication such as retinopathy (blindness), neuropathy (nerve damage), nephropathy (kidney disease), amputation, or skin ulcers. At the start of the enrollment period, there were 5,632 eligible persons—forty-one percent of all Fallon patients with Type 2 diabetes.

LifeMasters started recruiting beneficiaries in the summer of 1999. The company began by asking primary care physicians for permission to enroll their patients in the DM program. The physician acceptance rate was over 95 percent as the program rolled out. Eligible patients were then contacted by mail and telephone, receiving up to five calls. However, a substantial number of patients could not be reached because of inaccuracies in Fallon's telephone records. (We tried to obtain detailed data on telephone contacts, but found that reliable information could not be extracted from LifeMaster's phone logs.)

The process of obtaining physician permission, offering enrollment to eligible patients, and enrolling those who accepted took roughly six months. At the end of the initial six-month enrollment window, LifeMasters stopped actively recruiting patients in order to focus on program administration and the delivery of benefits to enrollees. In principle, patients remained eligible to enroll in the DM program. However, there was no effort by LifeMasters or Fallon to contact unenrolled patients or their physicians during the next five quarters.

Figure 1 shows the total number of enrolled patients during our sample period. The "first wave" of the admissions process began in June 1999, and LifeMasters enrolled 848 participants

(23 percent of all eligible patients). LifeMasters stopped soliciting patients in early 2000, and Figure 1 shows that passive enrollment was negligible over the next five quarters. In the second quarter of 2001, Fallon and LifeMasters resumed their recruiting activities in an effort to boost enrollment. Figure 1 shows that these efforts led to a steady increase in the number of beneficiaries.[6] However, they also changed the recruiting strategy in this this "second wave" of enrollment by focusing their efforts on eligible patients who had experienced deteriorating health outcomes during the intervening period. Because of this change, we focus on patients admitted during the initial recruiting wave below.

Once a patient enrolled in the DM program, they were immediately scheduled for a series of educational sessions. These lessons provided basic information about living with diabetes, and also explained how to use the resources available through the program's call center and web site. After completing these sessions, enrolled patients could call an automated phone system that would answer questions, record test results, or connect them to a nurse. They could also schedule regular one-on-one interactions with a case-worker.

Enrolled patients also received test kits and instructions for monitoring their own blood sugar levels. The kits were designed for use with the internet or telephone, so patients could easily send test results to LifeMasters. All of the information generated by enrolled patients was monitored and used in conjunction with claims data to generate "exception reports" that were screened and shared with physicians if there was a specific opportunity to intervene. Finally, LifeMasters periodically contacted enrolled patients to check on their health and satisfaction with the program.

4 Methods

Our primary goal is to identify the impact of the DM program on enrolled patients. Since we do not have a randomized-trial, we are forced to turn to non-experimental methods. In this section, we discuss our empirical strategy and methods of controlling for factors that might bias the results.

Our major concern is that patients who choose to enroll in the DM program may differ from those who decline, and that these differences could be correlated with observed outcomes. For example, individuals who are more likely to "take care of themselves," may also be more likely to sign up for DM, leading to a correlation between enrollment and health outcomes that simply reflects the impact of better (unobserved) self-care.

[6] During the second wave, LifeMasters also enrolled a number of patients who became eligible some time during the previous two years. We exclude these patients from the sample, as they were not eligible at the start of the initial enrollment period.

We control for patient heterogeneity by estimating difference in differences models.[7] This approach compares the change in outcomes for patients in the treatment group to the change in outcomes for patients in the control group. The treatment group difference controls for time-invariant patient and environmental characteristics that might be correlated with both enrollment and health. The control group difference captures any time-varying factors common to all patients. Another way to interpret this estimator is that the change in the control group serves as an estimate of the counterfactual—i.e. what would have happened to a treated patient in the absence of the intervention.

The basic difference-in-differences model can be specified as a two-way fixed effect linear regression:

$$y_{itk} = \alpha_k T_{ik} + \beta X_{it} + \gamma_i + \lambda_t + \varepsilon_{it} \tag{1}$$

where y_{itk} is an outcome for individual i who is in the k^{th} quarter of enrollment at time t, T_{ik} is an indicator variable that takes the value of 1 if individual i is in their k^{th} enrollment quarter and 0 otherwise, X_{it} is a vector of control variables that vary across both individuals and time, γ_i is a fixed-effect unique to individual i, λ_t is a time effect common to all individuals in period t, and ε_{it} is an individual time-varying error distributed independently across patients and independently of all γ_i and λ_t (Chamberlain, 1984). However, in the estimation we cluster the estimates of the standard errors at the individual level to control for possible serial correlation. The parameter α_k is an estimate of the average impact of the disease management program in the k^{th} quarter of treatment.

The key assumption in this model is that the change in the control group is an unbiased estimate of the counter-factual—i.e. the change the treatment group would have experienced had they not enrolled in the DM program. While it is not possible to test this assumption directly, we can use data from the pre-enrollment period to search for supporting evidence. In particular, we test the null hypothesis that pre-intervention trends for treated and control patients are identical. We perform this test by examining the α_k coefficients for $k < 0$ (where $k = 0$ corresponds to the quarter when a patient enrolled in the DM program). If these coefficients are not significantly different from zero, we should feel more comfortable with the exogeneity assumption (Heckman and Hotz 1989).

Our estimation sample is limited to patients who were eligible and offered the program in Spring of 1999. However, since Fallon's DM program had two waves of active recruitment, we have some flexibility in defining the treatment and control groups. For a treatment group, we

[7] The difference in difference estimator is one of the most widely used in the evaluation literature. See, among others, Angrist (1985), and Heckman, LaLonde and Smith (2000).

focus on patients who enrolled during the first recruitment wave. This is because LifeMasters offered the program to all eligible beneficiaries at the same time, and Fallon's poor phone records likely created a substantial amount of random variation in offers to enroll.[8] We exclude second wave enrollees from the treatment group as they were targeted in part based on recent health shocks. If the timing of enrollment were exogenous, we might consider using these patients as a second treatment group (i.e. estimate a treatment effect using data for only treated patients). However, this approach will produce biased estimates if there are time-varying factors correlated with both health and enrollment. In particular, any reversion to the mean in health outcomes would lead us to overstate the program's causal impact (Ashenfelter 1978).[9]

The presence of the second recruitment wave also leads to two possible control samples. Both control groups contain all eligible patients who did not enroll in the program during the first wave of recruitment—including the second wave enrollees. However, the first control group ends prior to the start of the second wave in 2001 (see Figure 1). The second control sample uses data from all available time periods, but discards any observations from second-wave enrollees once they join the DM program.

Though our models include patient fixed-effects to control for unobserved heterogeneity, one might still be concerned that the impact of treatment is not homogenous, but rather varies as a function of individual characteristics. We explore the issue of response heterogeneity by splitting the sample into patients with high and low "self-control" (where self-control is a measure of baseline CPG compliance described below) and estimating separate treatment effects for the two groups. We find that the largest beneficiaries of the DM program are patients with low self-control.

Another potential concern is attrition. If we focus on the period before the start of the second enrollment wave, 13.1 percent of all eligible patients exit our sample (6.9 percent through death and 6.2 percent through attrition). That number increases to twenty-seven percent if we extend the sample frame to the second quarter of 2002, with a mortality rate of 15 percent and 12.5 leaving Fallon for other reasons. This might pose a problem if, for example, patients who are more likely to die are also less likely to enroll in the DM program. As a robustness check, Table B-2 reproduces our main results using a balanced panel that excludes any patient who dies or exits the sample prematurely. The results are very similar those presented below.

Finally, we considered estimating a matched difference in difference model to further test the robustness of the results. However, matching estimators are very sensitive to the fit of a

[8] While there was considerable noise in the first wave of admissions, we do not assume exogenous assignment, since patients who were contacted could still decline the offer to enroll, and we cannot clearly identify who was or was not contacted. Discussions with LifeMasters suggest that they reached approximately half of all eligible patients during the first enrollment wave. Thus, a rough estimate of $Pr[enroll|contact]$ is 46 percent.

[9] We present estimates based on this approach as a robustness check in Tables B-3 and B-4.

first-stage enrollment model. We examined a variety of different enrollment models and found their explanatory power to be uniformly weak. As a result, we have refrained from estimating the matched difference in difference specification.

5 Data and Descriptive Statistics

Our primary data source is Fallon's patient medical claims and lab test results files. We use these data to measure compliance with clinical practice guidelines, health outcomes, and costs. We also know the patient's age, sex, and mailing address. Several additional demographic variables (e.g. race and income) were created by linking the patients' zip codes to data from the U.S. Census. For enrolled patients, we have some additional information provided during the registration process, including marital status and self-reported health behaviors, such as smoking and exercise. Finally, our mortality variable was constructed using an Internet site that provides social security death records.[10] Detailed information on variable definitions and data set construction are provided in Appendix A.

Table 1 reports sample means and standard deviations for all patients during the baseline year, i.e. the third quarter of 1998 through the second quarter of 1999. We organize the variables into four groups—patient demographics, CPG compliance, health outcomes, and costs. The first four columns in Table 1 provide statistics for all patients diagnosed with Type 2 diabetes, whom we divide into a low-risk (ineligible) and high-risk (eligible) population. The high-risk group accounts for 41 percent of all Fallon beneficiaries diagnosed with Type 2 diabetes. Not surprisingly, these patients are relatively older and sicker, and they spend roughly twice as much on medical care as the low-risk population. The last four columns in Table 1 report the same information for the high and low self-control sub-samples.

Focusing on the eligible population, we see that almost three-quarters of the patients in our sample received an HbA1c test during the pre-intervention year. Compliance on eye exams and cholesterol screening was roughly 50 percent, while only eight percent received a kidney exam. Conditional on testing, the mean HbA1c score was 8.34 (recall that 7.0 or less is considered "good control"), and twenty-one percent of the HbA1c scores were above 9.5 (the threshold for "poor control"). The quarterly inpatient admission rate was 8 percent, and 27 percent of eligible patients had a cardiac co-morbidity.[11] Finally, the average claims per patient-quarter were $2,260. While outpatient claims were $389 more than inpatient claims on average, the

[10] Specifically, we used a Perl script to input individual social security numbers into http://ssdi.genealogy.com/.

[11] We define a patient as co-morbid if they were diagnosed with Coronary Artery Disease (CAD), Chronic Obstructive Pulmonary Disease (COPD), or Congestive Heart Failure (CHF) prior to becoming eligible for enrollment.

inpatient claims display substantially more variation.

Table 2 compares sample means for treatment and control patients during the baseline year (1998 Q3 through 1999 Q2). The first three columns show baseline-year means for all eligible patients, along with P-values from a univariate T-test. The last two columns compare the treatment group in column one to our second control sample (patients who never enrolled). The treatment group and both control samples have very similar baseline demographics—only the Age and Self-control variables have a P-value less than 0.05. Our definition of self-control is based on receiving two or more of the recommended screening tests during the baseline year. We examine these tests in the second panel of Table 2, and find that treated patients are uniformly more likely to be tested (and therefore have higher self-control). While this might raise concerns about self-selection into the program, we condition on baseline compliance behaviors by splitting the sample and presenting separate results for the high and low self control groups.

The last two panels in Table 2 show that first wave enrollees and control patients have similar baseline health status. In particular, we cannot reject the null hypothesis that these groups have the same co-morbidity rate, inpatient admission rate and medical expenditures. While there is a significant difference in HbA1c levels, this may reflect the fact that treated patients were more likely to be tested.

5.1 Self-control

We expect patients who already take care of themselves to exhibit a smaller response to the DM program than those with less healthy pre-intervention behaviors. Our hypothesis is based on the belief that patient initiative is a key driver of compliance outcomes. This is almost certainly the case for diabetes, where the CPGs require patients to go to laboratories for recommended diagnostic tests, take their medicine as prescribed, exercise regularly and adhere to a strict diet.

To test this hypothesis we develop a measure of "self-control" based on compliance with the HEDIS guidelines published by National Committee for Quality Assurance (NCQA 2005).[12] The HEDIS guidelines suggest that persons with diabetes should have a suite of annual diagnostic tests including an HbA1c exam, a retinal exam, a lipids (cholesterol) panel, and a test for kidney function. According to our definition, a patient has "high self-control" if they received two or more of the four recommended screening tests (HbA1c, Cholesterol, Eyes, and Kidney) during the baseline year and "low self-control" if they received one or fewer of these tests. By

[12] In the behavioral economics literature "self-control" is commonly associated with the shape of an individual's time preferences. While many of the papers in this literature focus on savings and investment decisions, there are several studies—such as Fuchs (1982) or DellaVigna and Passerman (2005)—that use health-related behaviors as a measure of individual self-control.

this definition, 65 percent of the eligible population had high self-control, while 35 percent had low self-control.[13]

Table 3 presents a pairwise correlation matrix for the four baseline compliance indicators used to construct our measure of self-control. All of the correlations are positive, which is consistent with the "standard model" of self-care as an investment. However, the cross-sectional correlations are quite small—five of the six coefficients are less than 0.15—suggesting that there is no single individual-specific factor, such as discount rates or physician quality, that explains most of the variation in CPG compliance. The small correlation coefficients also reflect the fact that these tests are not generally performed by the same type of medical professional. In particular, they cannot generally be combined into a single physician visit.

To examine whether our measure is a good proxy for other healthy behaviors, we created a compliance index (i.e. a count of the recommended tests received by each patient during the baseline year, ranging from zero to four) and regressed this variable on several standard proxies for self-control, including smoking, weight control, and regular exercise. We were only able to conduct this exercise for the treatment group, because these data were collected as part of the DM program's enrollment process, and hence are not available for the controls. Table 4 presents the results.

The first column in Table 4 shows that two of our three behavioral proxies for self control are correlated with the baseline compliance index at statistically significant levels. In particular, smokers receive about 13 percent less preventive testing than non-smokers, while patients who exercise on a regular basis receive almost 8 percent more. This suggests that our index of diagnostic test compliance will also reflect compliance with unmeasured CPGs, such as diet and exercise. We also find that Age and Medicare enrollment are positively correlated with baseline self-control. There are a variety of reasons why older patients might take better care of themselves (e.g. retirement), and it is not surprising that changes in medical coverage explain some of the variation in compliance.

While the four tests in our compliance index require a certain amount of patient initiative (e.g. several trips to different labs and physicians) it is possible that this measure reflects physicians' failure to prescribe the appropriate tests rather than individual self-control. We check whether variation in the compliance index reflects physician as opposed to individual behaviors by including physician fixed effects, and those results are reported in the second and third column of Table 4. The results continue to indicate a significant correlation between aggregate CPG compliance and behavioral measures of self-control even after controlling for

[13] Our composite measure will mask some of the variation in compliance behavior for individual tests. However, we have reproduced all of our primary results using alternative definitions of self-control that place more weight on specific tests (e.g. HbA1c screening).

patient age, health coverage, and physician heterogeneity. In fact, adding physician effects only explains an additional 1% of the total variation in the compliance index, suggesting that most of this variation is driven by patient rather than physician behaviors.

Finally, returning to the last four columns in Table 1, we find several differences in the baseline-year demographics of the high and low self-control sub-samples. On average, high self-control patients are older, sicker, and more costly. The average age of a patient with low self-control is 63.5 years, while a patient with high self-control averages 66.7 years. This helps to explain the 13 percent difference in Medicare enrollment across these two groups. While the mean HbA1c score of low self-control patients is somewhat higher, this measure should be treated with caution because of the selection bias introduced by our definition of self-control. In particular, only 34 percent of the low self-control patients received an HbA1c exam during the baseline year, compared to 94 percent of those with high self-control. There is no difference in the baseline inpatient admission rate of the two groups, but high self-control patients are 5 percent more likely to have a serious cardiac co-morbidity. Finally, the quarterly claims rate was $388 greater for patients in the high self-control group—which is not particularly surprising given the differences in age and co-morbidity status. On most other dimensions, the high and low self control samples appear to be quite similar.

6 Results

Our main results are presented in three tables. Table 5 emphasizes the timing of the treatment effects. In particular, we show that treated and control patients have similar pre-intervention trends, and that the impact of the DM program is apparent within 2 or 3 quarters of enrollment. These results provide support for our identification strategy. In Tables 6 and 7, we adopt a more parsimonious specification (i.e. a single treatment effect parameter) and examine a broader set of outcomes.

6.1 Timing

In Table 5 we focus on three dependent variables that capture changes in compliance behavior, health outcomes and the total cost of care. The first of these variables is the compliance index used in Table 4. The second variable is an indicator for "poor control" over blood sugar levels, as indicated by a score greater than 9.5 on an HbA1c test.[14] The third variable is the log of

[14] We are somewhat cautious with this variable, since we only observe an HbA1c score when a patient is tested. As a result, we have relatively few pre-intervention scores for the low self-control patients. Moreover, since poor health leads to more frequent testing, we have more scores for less healthy patients. We use an indicator variable (rather than a continuous test score) to alleviate any bias introduced by sampling on past test results. Since

total expenditures, which we define as the sum of inpatient, outpatient and pharmaceutical claims. (We add $1 to every patient quarter, so that the smallest value taken by our log cost measure is zero.)

For each dependent variable, we estimate the difference in difference specification in Equation (1). We condition on baseline CPG compliance by splitting the sample into groups with low and high self-control, and we allow the impact of the DM program to vary over time by including a full set of T_{ik}—where k indexes a patient's cumulative exposure to the DM program (omitting the dummy variable corresponding to $k < -2$). For these regressions, we include all eligible patients and limit the sample frame to the time-period before the first quarter of 2001 to avoid concerns related to the second-wave of enrollment.[15] The first three columns in Table 5 present results for patients with low self-control, while columns 4 through 6 present results for the high self-control group.

The first question addressed by Table 5 is related to identification: do the treatment and control patients have different pre-intervention trends? The first column of results suggests that treated patients with low self-control do show a small increase in CPG compliance (relative to unenrolled patients with low self-control) prior to enrollment. However, this effect doubles at enrollment and grows in size and significance over subsequent quarters. We find no evidence of a difference in pre-treatment trends for HbA1c or total cost in the low self-control sub-sample. Moreover, there is a sharp break in the secular trend for each of these variables following enrollment (in quarter 1 for costs and quarter 2 for HbA1c scores). The sudden change in all three outcomes within one or two quarters of enrollment provides evidence to support the validity of our difference in differences estimator.

Turning to the high self-control patients, we find no evidence of a difference in the pre-treatment trends for compliance or HbA1c. While we do not observe a clear break in the compliance trend, there is a pronounced improvement in HbA1c scores during the first and second enrollment quarter. Finally, the last column in Table 5 provides some evidence that total costs are declining for the high self-control patients prior to enrolling in the DM program. However, this trend is not particularly strong, as relative costs rise in quarters 1 and 2 before dropping again in the third enrollment quarter.

Overall, the pre-treatment coefficients in Table 5 suggest very small differences between the treated and control patients before the DM intervention. In fact, it is remarkable that only two of the twelve pre-treatment coefficients reach the 5 percent significance level given that

our diff-in-diffs estimator is identified by changes, this essentially throws out variation in HbA1c levels except for patients who cross the critical threshold.

[15] The "pre treatment" dummies T_{ik}, where $k < 0$ are also set to zero for patients who enroll during the second wave.

our sample size ranges from two to three thousand patients. These results support LifeMasters contention that a number of plausibly exogenous factors—particularly the lack of good phone records—had a significant impact on the first wave enrollment process.[16]

The second question addressed by Table 5 is, "How fast does the DM program work?" To answer this question, we focus on the coefficients measuring differences between enrolled and unenrolled patients in the first four quarters of enrollment. Our estimates suggest that the response is remarkably fast for patients in the low self-control group.

By the second quarter of enrollment, low self-control patients exhibit a compliance increase of 0.21 tests, "poor" HbA1c scores fall by 20 percent and total costs decline by 65 percent relative to the unenrolled patients. All of these improvements persist. For patients with high self-control, the benefits are smaller and take longer to materialize. For example, by the fourth enrollment quarter compliance increases by 0.11 tests per patient, HbA1c control improves by 14 percent, and costs fall by 10 percent (though the last effect is not statistically significant). These differences in the timing and magnitude of the estimated treatment effects suggest that DM interventions are more effective for patients who lack the initiative to comply with CPG guidelines.

As an additional robustness check, we estimate a series of models that include a linear time-trend for treated patients and present the results in Appendix Table B-1.[17] For the low self-control patients, this has little impact on the HbA1c and log Total Cost estimates — the time-trend is statistically insignificant and post-intervention coefficients are qualitatively similar. Introducing a time trend does eliminate post-intervention changes in the compliance index for patients with low self-control (though this may not be surprising given the strong upward trend in the post-intervention coefficients). For the high self-control group, the compliance and HbA1c results are qualitatively similar, while the financial impacts are eliminated.

Overall, we interpret the results in Table 5 and B-1 as evidence that patients who enrolled in the DM program during the first wave of admissions are not systematically different from those who did not enroll, but experienced sharp changes in compliance, health and financial outcomes following admission to the program. This suggests that our differences in differences identification strategy is valid, and that the results of the intervention appeared rather quickly. There is also evidence that the program's impact was greater for patients with low self control.

[16]Table B-3 in the appendix replicates this analysis using second-wave enrollees as the control sample (i.e. dropping all patients who never enroll). We find significant differences in the results—particularly for compliance and costs in the low self-control group—which suggest that many late-enrollees were admitted to the DM program following adverse health shocks.

[17]In this specification, the post-intervention coefficients are identified by differences between actual outcomes and an extrapolation based on the pre-intervention time-trend. We restrict attention to the first four post-enrollment quarters, to avoid gross exaggeration of the treatment effects based on lengthy extrapolations.

6.2 Additional Outcomes

In Tables 6 and 7 we expand the estimation sample, adopt a more parsimonious specification (with a single treatment effect) and examine a broader set of outcome measures. Specifically, we use data extending through June 2002, and estimate a single treatment effect parameter (i.e. we restrict $\alpha_k = 0$ for $k < 0$ and $\alpha_k = \alpha$ for $k \geq 0$). Table 6 examines compliance and health outcomes, including mortality and morbidity. Table 7 focuses on financial impacts by breaking total expenditures into inpatient, outpatient and pharmaceutical claims. In both tables, we estimate separate models for patients with high and low self-control, and we consider two different control samples: eligible patients who never enroll (excluding second-wave enrollees), and eligible patients who are not currently enrolled (dropping second-wave enrollees after they enter the program).[18]

We begin by examining the DM program's long-run impact on compliance and health outcomes. The first column in Table 6 shows that the program had a significant impact on compliance for both the high and low self control groups. Beneath the standard errors we express each coefficient in terms of a percentage change relative to the baseline year. Not surprisingly, both absolute and percentage improvements in compliance are greater for the low self control patients (who, by definition, had lower baseline compliance levels).

The second column in Table 6 shows significant improvements in the upper tail of the HbA1c distribution following admission to the DM program. In particular, the probability that a patient has poor control over their blood sugar falls by 20 to 40 percent. While the point estimates are slightly larger for patients with high self control, we cannot reject the hypothesis that the effects are the same in the low self control group.

In the third column, the dependent variable is a dummy for inpatient admissions. The point estimates suggest that there is a 2 percentage point decline in the inpatient admission rate for the low self-control group, and a one percentage point decline for the high self-control patients. While these are large effects, equivalent to a 10 or 25 percent drop in the baseline inpatient admission rate, they are not statistically significant.

The last two columns in Table 6 examine morbidity and mortality. Morbidity is an indicator variable that equals one in the quarter when a patient is diagnosed with a serious complication that is clinically linked to diabetes. Specifically, morbidity indicates Coronary Artery Disease, Congestive Heart Failure, Coronary Obstructive Pulmonary Disorder, Stroke, Blindness, Kidney Failure, or Amputation. Mortality is a dummy variable indicating that the patient died in a given quarter.[19]

[18] Table B-4 presents comparable estimates for the treated patients only (i.e. using the pre-enrollment period to estimate counterfactual outcomes).

[19] We do not provide mortality estimates using late enrollees as the control sample, since it is impossible for a

Since neither of these variables is a repeated outcome, we cannot estimate a difference in differences model containing individual fixed-effects. Instead, we estimate a discrete-time hazard model. In particular, we assume that the patient-specific hazard rate (i.e. the quarterly probability of morbidity or mortality) takes a logistic form $h_{it} = \exp\{X_{it}\beta + \lambda_t\}$. In this expression, the λ_t are quarterly fixed-effects that provide for a flexible parameterization of the baseline hazard, while the X_{it} are a set of patient-quarter covariates that may or may not change over time. This model can be estimated as a simple logit regression (Jenkins, 1995). We are interested in the coefficient β on a time-varying indicator for enrollment in the DM program. In Table 6, we present marginal effects (calculated at the sample means) and robust standard errors. A full set of estimates for these models (which include a large number of additional controls) are presented in Tables B-5 and B-6.

The fourth column in Table 6 shows that there is a 0.8 percentage point drop in the quarterly incidence of serious complications for patients with low self-control. This effect is statistically significant at the ten percent level in our first control sample (5 percent for the second set of controls) and corresponds to a 26 percent reduction in the baseline rate at which patients develop serious complications. There is no morbidity effect in the high self-control group. The final column shows a small decline in mortality rates: 0.1 percentage points for patients with low self-control and 0.2 percentage points for the high self-control sample. The high self-control result corresponds to a 21 percent decline in the mortality rate, and is significant at the ten percent level for the first control sample; though both magnitude and significance decline if we include second wave enrollees in the control group.

Overall, Table 6 suggests that that the DM program produces improvements in compliance and health outcomes for all enrolled patients. The compliance and HbA1c effects are large and statistically significant. While the point estimates for inpatient admissions, morbidity and mortality generally imply large effects, they are not precisely estimated. However, before concluding that the DM program has little or no impact on these outcomes, we would like to collect more evidence. In particular, we might expect any impact on these outcomes to appear after a longer exposure to the DM program (e.g. because major health improvements are tied to short-run improvements in other measures, such as compliance and HbA1c).

Table 7 examines the financial impact of the DM program. Specifically, we separate the total cost variable into three parts: inpatient, outpatient and pharmaceutical claims. Since these variables enter in logs, we can interpret the coefficient estimates as a percentage change. Beneath the coefficient estimates, we present an estimate of the savings per patient-quarter in US dollars obtained from a separate regression run in levels. While these estimates are useful

patient to die before they enroll the DM program.

for thinking about the potential magnitude of the program's cost savings, we prefer the log specification which is less sensitive to the outliers frequently found in medical claims data.

The first column in Table 7 presents an estimate of total cost savings. These coefficients suggest that costs drop by 55 percent for the low self-control patients, and 5 percent for those with high self-control (the latter result is not statistically significant). Where do these cost savings come from? The second column examines inpatient costs. While the point estimates suggest a 17 percent reduction in inpatient claims for the low self-control patients and a 9 percent decline for those with high self-control, neither estimate is statistically significant. However, column three shows a large (roughly 41 percent) drop in outpatient costs for the low self-control group. The change in outpatient costs for the high self-control group was negligible. The last column in Table 7 shows a 44 percent decline in pharmaceutical spending for the low self-control group and a negligible drop for the high self-control patients.

Table 7 shows that DM can produce substantial financial benefits. However, these benefits appear to be concentrated within the low self-control population. This finding has significant implications for targeting DM interventions towards patient populations that are likely to yield a large financial return. It is also a reassuring outcome, since these are precisely the individuals we would expect to benefit most from a preventive intervention. It is also interesting to note that the results for inpatient costs, morbidity and mortality are economically meaningful but not statistically significant within our three year study window. Our analysis suggests that DM programs may produce significant short-run cost savings even if some of these larger benefits of DM take longer to materialize.

However, we remain cautious in our interpretation of the links between health and financial outcomes in these data. In particular, one plausible interpretation of these financial results is that patients in the low self-control group are using the DM program as a substitute for more frequent interaction with their primary-care physician (which would show up as an outpatient expenditure in our data). Since outpatient cost reductions are clearly driving the overall financial benefits, it would be interesting to examine this issue using more detailed data and perhaps a longer post-intervention time-window.

6.3 Financial Implications

To obtain a more complete picture of the financial benefits produced by a DM program, we should account for administrative costs and also patterns of enrollment and attrition. Unfortunately, we do not have any data on the costs of administering Fallon's DM program that can be compared to our previous estimates. However, discussions with LifeMasters indicate that if our results are accurate, the program would have produced substantial net savings. If we

use the larger of the two published cost estimates cited above ($117 per patient-year), simple calculations suggest that Fallon saved several hundred dollars per enrolled patient per year.

To simulate Fallon's *actual* cost savings, we can combine our regression results with data on patterns of enrollment and attrition. We assume that all of the financial benefits produced by the DM program came from patients with low self-control, and that these savings followed a step function—producing no benefits during the first six months of enrollment before jumping to $560 per enrolled patient per quarter thereafter.[20] Under these assumptions, simple calculations suggest that Fallon's DM program reduced total claims by $1.09 million over a three year period. Thus, our results indicate that DM can produce significant cost savings in the short- and medium-run.

Finally, it is important to point out that all of the financial benefits described above—including the large potential savings for the Medicare program described in the introduction—are presumably small in comparison to the welfare gains associated with improved quality of life for individuals living with diabetes. These welfare benefits will continue to grow as the prevalence of diabetes and other chronic illnesses increases due to an aging population.

7 Conclusions

The main finding of this paper is that Fallon's diabetes DM program had a significant impact on compliance behaviors, health outcomes, and the overall cost of care. Specifically, we find a significant increase in CPG compliance, a reduction in HbA1c levels, a reduction in the onset of serious complications, significant financial benefits and some evidence of reduced morbidity and mortality rates. We also find that most of the benefits—especially the financial ones—were realized by patients with low self-control, i.e. those who did not comply with clinical guidelines during the pre-intervention period.

To the extent that the results of this empirical case study generalize to the broader patient population, they have two important policy implications. First, DM programs may be a financially attractive way of improving the health of chronically ill patients, and a potential solution to the looming crisis in health spending. We find that there were relatively large net financial benefits associated with Fallon's diabetes DM program. Perhaps more importantly, these benefits appeared relatively quickly—within 6 months of enrollment. While this primarily

[20] This figure was computed by multiplying the baseline annual claims rate for a patient with low self-control (see Table 1) by 35 percent—the lower end of the 95 percent confidence interval for the treatment effect in this sub-sample. Alternatively, we could take the point estimate of -0.435 and calculate expected savings by exponentiating the predicted change in log total expenditures, which yields an estimate of $440 per enrolled patient per quarter, and total savings of $850,000 (for the only the low self-control patients) over the entire three-year period.

reflects reduced outpatient utilization, there is also some evidence of a decline in the inpatient admission rate.

Second, our findings suggest that large DM adopters, such as the Medicare program, would be well advised to target DM services to patients who are not complying with evidence-based CPGs in order to maximize the return on these investments. The non-compliant patients in our study experienced significantly more improvement upon enrolling in the DM program. These results confirm the validity of the evidence-based CPGs developed over the last decade to encourage the adoption of best practices in preventive care. At the same time, given the strong correlation between CPG compliance and a number of widely used proxies for "self-control," these results point to the economic significance of self-control problems and the potentially powerful role of behavioral interventions like DM.

In conclusion, this paper points towards a number of opportunities for future research. One promising avenue is to examine specific mechanisms through which DM might change individual patient behaviors. A second avenue is to examine how DM performs in other health care delivery environments, which might differ in terms of patient demographics, physician incentives, insurance schemes, or competition. Finally, much work remains to be done on the broader question of explaining why individuals do or do not take care of themselves.

References

ACCORD (2008). Effects of intensive glucose lowering in type 2 diabetes: The Action to Control Cardiovascular Risk in Diabetes group. *New England Journal of Medicine 358*(24), 2545–2559.

Angrist, J. (1985). Introduction to the JBES Syposium on Program and Policy Evaluation. *Journal of Business and Economic Statistics 13*, 249–88.

Ashenfelter, O. (1978). Estimating effect of training-programs on earnings. *Review of Economics and Statistics 60*(1), 47–57.

Beaulieu, N., D. Cutler, and K. Ho (2003). The business case for diabetes disease management at two managed care organizations: A case study of healthpartners and independent health association. Field report, The Commonwealth Fund.

Becker, G. and K. Murphy (1988). A theory of rational addiction. *Journal of Political Economy 96*(4), 675–700.

Berg, G. and S. Wadhwa (2007). Health services outcoems for a diabetes disease management program for the elderly. *Disease Management 10*(4), 226–234.

Bray, K., R. Turpin, K. Jungkind, and G. Heuser (2008). Defining success in diabetes disease management: Digging deeper in the data. *Disease Management 11*(2), 119–128.

Casalino, L., R. R. Gillies, S. M. Shortell, J. A. Schmittdiel, T. Bodenheimer, J. C. Robinson, T. Rundall, N. Oswald, H. Schauffler, and M. C. Wang (2003). External incentives, information technology, and organized processes to improve health care quality for patients with chronic diseases. *JAMA-Journal of the American Medical Association 289*(4), 434–441.

CBO (2004). An anlaysis of the literature on disease management programs. *http://www.cbo.gov/ftpdocs/59xx/doc5909/10-13-DiseaseMngmnt.pdf Accessed June 2008.*

Chamberlain, G. (1984). Panel data. In Z. Griliches and M. Intrilligator (Eds.), *Handbook of Econometrics, vol. 2*. North-Holland.

Chandra, A., J. Gruber, and R. McKnight (2007). Patient cost-sharing, hospitalization offsets, and the design of optimal health insurance for the elderly. NBER Working Paper No. 12972.

Cutler, D. and E. Glaeser (2005). What explains differences in smoking, drinking, and other health-related behaviors? NBER Working Paper No. 11100.

DellaVigna, S. and M. Paserman (2005). Job search and impatience. *Journal of Labor Economics 23*(3), 527–587.

Felt-Lisk, S. and G. P. Mays (2002). From the field - back to the drawing board: New directions in health plans' care management strategies. *Health Affairs 21*(5), 210–217.

Fuchs, V. (1982). Time preference and health : an exploratory study. In V. R. Fuchs (Ed.), *Economic Aspects of Health*, pp. 93–120. Chicago, IL: University of Chicago Press.

Gillespie, J. and L. Rossiter (2003). Medicaid disease management programs. *Disease Management health Outcomes 11*(6), 345–361.

Goldman, D. and J. Smith (2002). Can patient self-management help explain the ses health gradient? *Proceedings of the National Academy of Science 99*(16), 10929–34.

Grossman, M. (1972). On the concept of health capital and the demand for health. *Jounral of Political Economy 80*(2), 223–55.

Heckman, J. and V. Hotz (1989). Choosing among alternative nonexperimental methods for estimating the impact of social programs: The case of manpower training. *Journal of the American Statistical Association 84*(408), 862–874.

Heckman, J., H. Ichimura, and P. Todd (1998). Matching as an econometric evaluation estimator. *Review of Economic Studies 65*(2), 261–194.

Heckman, J., R. LaLonde, and J. Smith (2000). The economics and econometrics of active labor market programs. In O. Ashenfelter and D. Card (Eds.), *The Handbook of Labor Economics, Vol. 3A*. North-Holland.

Hoffman, C. and D. Rice (1995). Estimates based on the national medical expenditure survey. *UCSF Institute for Health and Aging*.

Hogan, P., T. Dall, and P. Nikolov (2003). Economic costs of diabetes in the u.s. in 2002. *Diabetes Care 26*(3), 917–932.

IOM (2001). Crossing the quality chasm: A new health system for the 21st century. Technical report, Institute of Medicine,.

Jenkins, S. (1995). Easy estimation methods for discrete-time duration models. *Oxford Bulletin of Economics and Statistics 57*(1), 129–137.

Knight, K., E. Badambarav, J. Henning, V. Hasselblad, A. Gano, J. Ofman, and S. Weingarten (2005). A systematic review of diabetes disease management programs. *The American Journal of Managed Care 11*(5), 242–250.

Laibson, D. (1997). Golden eggs and hyperbollic discounting. *Quarterly Journal of Economics 112*, 443–477.

Lairson, D., S. Yoon, P. Carter, A. Greisinger, K. Talluri, M. Aggarwal, and O. Wehmanen (2008). Economic evaluation of an intensified disease management system for patients with type 2 diabetes. *Disease Management 11*(2), 79–94.

Levit, K., C. Smith, C. Cowan, C. Sensenig, A. Catlin, and e. al (2004). Health spending rebound continues in 2002. *Health Affairs 23*(1), 147–159.

McGinness, J. and W. Foege (1993). Actual causes of death in the united states. *Journal of the American Medical Association 270*(18), 2207–12.

McWilliams, J., E. Meara, Z. A., and J. Ayanian (2007). Use of health services by previously uninsured medicare beneficiaries. *New England Journal of Medicine 357*, 143–153.

Medicare (2005). A data book: Healthcare spending and the medicare program. Technical report.

NCQA (2005). The state of health care quality 2005. Technical report, National Committee for Quality Assurance.

O'Donohugh, T. and M. Rabin (1999). Doing it now or later? *American Economic Review 89*(1), 103–124.

Rosenbaum, P. and D. Rubin (1983). The central role of the propoensity score in observational studies for causal effects. *Biometrika 70*, 41–55.

Rothman, R., D. DeWalt, R. Malone, B. Bryant, A. Shintani, B. Crigler, M. Weinberger, and M. Pignone (2004). Influence of patient literacy on the effectiveness of a primary carebased diabetes disease management program. *Journal of the American Medical Association 292*(14), 1711–1716.

Sturm, R., J. Ringel, and T. Andreyeva (2004). Increasing obesity rates and disability trends. *Health Affairs 23*(2), 199–205.

Wennberg, J. and A. Gittlesohn (1973). Small area variations in health care delivery. *Science 182*, 1102–8.

Table 1: Pre-intervention Baseline Summary Statistics*

	Ineligible (Low Risk) Diabetes Patients		Eligible (High Risk) Diabetes Patients		Eligible High Self-control		Eligible Low Self-control	
Variable	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
	Demographics							
Age	60.86	16.11	65.58	13.75	66.72	12.23	63.47	15.97
Male	0.51	0.50	0.53	0.50	0.53	0.50	0.53	0.50
Urban (Zip)	0.85	0.23	0.84	0.23	0.84	0.23	0.84	0.23
White (Zip)	0.89	0.11	0.90	0.09	0.90	0.09	0.89	0.09
Income (Zip)	49,670	13,980	49,279	13,421	49,580	13,393	48,724	13,461
Medicare	0.46	0.50	0.59	0.49	0.64	0.48	0.51	0.50
Medicaid	0.04	0.19	0.03	0.16	0.02	0.14	0.04	0.19
	Compliance Behavior							
HbA1c Compliant	0.47	0.50	0.73	0.44	0.94	0.23	0.34	0.47
Eye Compliant	0.51	0.50	0.56	0.50	0.73	0.44	0.25	0.43
Lipid Compliant	0.45	0.50	0.48	0.50	0.70	0.46	0.09	0.28
Kidney Compliant	0.03	0.18	0.08	0.27	0.12	0.33	0.00	0.04
Compliance Index	1.46	1.05	1.86	1.03	2.50	0.60	0.68	0.47
	Health Status							
HbA1c Level	7.19	1.37	8.34	1.85	8.26	1.79	8.72	2.06
HbA1c $\geq$ 9.5	0.06	0.24	0.21	0.40	0.19	0.39	0.29	0.45
Inpatient Admit	0.03	0.18	0.08	0.27	0.08	0.27	0.08	0.27
Cardiac co-morbidity	0.12	0.33	0.27	0.44	0.29	0.45	0.24	0.43
	Cost of Care							
Total Claims	$1,161	3,442	$2,260	6,815	$2,395	7,107	$2,007	6,218
Inpatient Claims	$251	2,274	$733	5,201	$750	5,514	$703	4,554
Outpatient Claims	$665	1,658	$1,122	2,410	$1,197	2,490	$981	2,246
Pharma Claims	$245	340	$405	431	$448	427	$324	428
Total Patients	8,053		5,632		3,653		1,979	

*Baseline year is 1998Q3 through 1999Q2.

Table 2: Pre-intervention Baseline Means: Treatment vs. Control

	All Eligible Patients			Never Enrolled	
	Treated*	Controls 1*	P-value†	Controls 2*	P-value†
	Demographics				
Age	64.58	65.75	0.01	66.05	0.00
Male	0.51	0.54	0.17	0.55	0.07
Medicare	0.57	0.60	0.21	0.61	0.10
Medicaid	0.02	0.03	0.06	0.03	0.07
Urban	0.84	0.84	0.92	0.84	0.86
White	0.90	0.89	0.13	0.89	0.08
Income	49,448	49,249	0.68	49,173	0.58
Attrition	0.12	0.13	0.88	0.14	0.28
High Self-control	0.78	0.63	0.00	0.63	0.00
	Compliance Behavior				
HbA1c Exams	2.03	1.42	0.00	1.42	0.00
Eye Exam	1.16	0.99	0.00	1.01	0.00
Lipids Panel	0.89	0.71	0.00	0.69	0.00
Kidney Exam	0.13	0.08	0.00	0.08	0.00
Compliance i	2.21	1.80	0.00	1.80	0.00
	Health Status				
HbA1c Level	8.91	8.21	0.00	8.23	0.00
HbA1c > 9.5	0.20	0.23	0.00	0.23	0.00
Comorbidity	0.30	0.27	0.08	0.27	0.17
Inpatient Visit	0.25	0.22	0.06	0.23	0.15
	Cost of Care				
Total Cost	7,538	7,192	0.49	7,425	0.82
Inpatient Claims	2,776	2,879	0.75	3,007	0.49
Pharma Claims	2,091	1,491	0.00	1,488	0.00
Total Patients	848	4,784		4,118	

*Treated patients are those who enrolled before 2000Q2 (see Figure 1). Controls 1 include all eligible patients not in the treatment group. Controls 2 include only patients who never enroll in the DM program. †Based on two sample T-test with unequal variance.

Table 3: Baseline CPG Compliance Correlations*

	HbA1c	Lipids	Retinal	Kidney
HbA1c Exam	1.00			
Cholesterol Exam	0.34	1.00		
Retinal Exam	0.08	0.06	1.00	
Kidney Exam	0.14	0.13	0.04	1.00

*Correlations among four dummy variables indicating that a patient received a recommended screening exam during the baseline year. Sample is 5,632 eligible patients.

Table 4: Baseline Compliance and Self-Control

	OLS Regression DV = Compliance Index		
Smoking	-0.275 (0.091)**	-0.216 (0.094)*	-0.198 (0.099)*
Exercise	0.164 (0.055)**	0.194 (0.058)**	0.181 (0.062)**
Overweight	-0.066 (0.086)	-0.085 (0.084)	-0.126 (0.089)
Married	0.140 (0.067)*	0.083 (0.068)	0.050 (0.072)
log (Age)	0.551 (0.176)**	0.513 (0.192)**	0.494 (0.204)*
Male	0.010 (0.052)	-0.000 (0.057)	-0.040 (0.060)
Medicare	0.221 (0.079)**	0.271 (0.081)**	0.234 (0.085)**
Medicaid	0.188 (0.192)	0.227 (0.190)	0.229 (0.200)
Comorbid	0.093 (0.057)	0.102 (0.063)	0.109 (0.066)
log (Income)	0.080 (0.096)	-0.091 (0.115)	0.052 (1.016)
Constant	-2.236 (1.415)	-0.189 (1.633)	-2.189 (10.602)
Patients	1,508	1,508	1,508
R-squared	0.06	0.07	0.19
Physician FEs (F-Test)	N	216 1.43**	216 1.25*
Zip Code FEs (F-Test)	N	N	151 1.13

Robust standard errors in parentheses. *Significant at 5%; **significant at 1%. Estimation sample is 1,508 eligible patients who enrolled in the DM program.

Table 5: DM Impact by Enrollment Quarter: 1998Q3 to 2001Q1

	Low Self-Control†			High Self-Control		
	Compliance Index	Hba1c >9.5	log Total Costs	Compliance Index	Hba1c >9.5	log Total Costs
Quarter = -2	0.072 (0.035)*	-0.117 (0.066)	0.119 (0.087)	0.001 (0.029)	-0.043 (0.024)	-0.042 (0.052)
Quarter = -1	0.096 (0.054)	-0.075 (0.063)	0.000 (0.120)	-0.057 (0.037)	-0.030 (0.025)	-0.140 (0.054)*
Enrollment Quarter	0.183 (0.064)**	-0.017 (0.056)	-0.569 (0.156)**	0.001 (0.040)	-0.086 (0.023)**	-0.067 (0.048)
Quarter = +1	0.214 (0.076)**	-0.203 (0.061)**	-0.650 (0.175)**	0.074 (0.044)	-0.140 (0.023)**	-0.043 (0.052)
Quarter = +2	0.215 (0.081)**	-0.190 (0.058)**	-0.756 (0.181)**	0.093 (0.044)*	-0.124 (0.024)**	-0.119 (0.057)*
Quarter = +3	0.158 (0.083)	-0.230 (0.071)**	-0.675 (0.163)**	0.110 (0.045)*	-0.136 (0.028)**	-0.100 (0.063)
Quarter 4+	0.309 (0.084)**	-0.165 (0.057)**	-0.336 (0.115)**	0.046 (0.046)	-0.113 (0.022)**	-0.150 (0.059)*
Patient fixed-effects	Y	Y	Y	Y	Y	Y
Calendar-quarter FEs	Y	Y	Y	Y	Y	Y
Patient-quarters	20,476	4,105	20,476	38,311	14,819	38,311
Patients	1,979	1,281	1,979	3,653	3,011	3,653

Notes: OLS regression. Robust standard errors clustered by patient in parentheses; *significant at 5%; **significant at 1%. HbA1C models have fewer observations because tests do not occur quarterly.
†Low self-control patients have a baseline compliance-index less than two.

Table 6: Compliance and Health Impacts: 1998Q3 to 2002Q2

	Compliance Index	HbA1c ≥ 9.5	Inpatient Visit	Morbidity†	Mortality†
	Sample = First wave and unenrolled patients				
Low Self-control†† (1,714 Patients)	0.268 (0.055)**	-0.083 (0.041)*	-0.021 (0.012)	-0.008 (0.005)	-0.001 (0.002)
	51.29%	-27.43%	-26.05%	-25.78%	-16.20%
High Self-control (3,252 Patients)	0.077 (0.032)*	-0.097 (0.016)**	-0.011 (0.008)	0.002 (0.003)	-0.002 (0.001)
	4.14%	-48.77%	-13.10%	6.26%	-21.54%
	Sample = All patients (dropping second wave at enrollment)				
Low Self-control (1,979 Patients)	0.232 (0.054)**	-0.065 (0.027)*	-0.019 (0.012)	-0.009 (0.005)*	-0.000 (0.001)
	44.72%	-19.79%	-23.36%	-27.22%	-7.17%
High Self-control (3,653 Patients)	0.066 (0.032)*	-0.076 (0.016)**	-0.010 (0.008)	0.002 (0.003)	-0.001 (0.001)
	3.50%	-31.83%	-12.21%	4.85%	-13.92%

Notes: Each coefficient is estimated in a separate OLS regression with patient and calendar-quarter fixed-effects. Robust standard errors clustered by patient in parentheses; *significant at 5%; **significant at 1%. Percentage changes are measured relative to baseline-year means. †Marginal effects and robust standard errors from a discrete-time logistic hazard model; see text for a discussion of the specification and Tables B-5 and B-6 for a complete set of results. ††Low self-control patients have a baseline compliance-index less than two.

Table 7: Financial Impacts: 1998Q3 to 2002Q2

	log Total Claims	log Inpatient Claims	log Outpatient Claims	Log Pharma Claims
	Sample = First wave and unenrolled patients			
Low Self-control†	-0.557	-0.170	-0.411	-0.435
(1,714 Patients)	(0.094)**	(0.107)	(0.117)**	(0.098)**
	($238)	($89)	($114)	($36)
High Self-control	-0.046	-0.094	-0.009	0.047
(3,252 Patients)	(0.041)	(0.070)	(0.056)	(0.053)
	($201)	($235)	+$23	+$11
	Sample = All patients (dropping second wave at enrollment)			
Low Self-control	-0.577	-0.147	-0.436	-0.487
(1,979 Patients)	(0.093)**	(0.106)	(0.116)**	(0.096)**
	($207)	($62)	($105)	($40)
High Self-control	-0.058	-0.082	-0.025	0.018
(3,653 Patients)	(0.040)	(0.069)	(0.055)	(0.052)
	($162)	($185)	+$17	+$6

Notes: Each coefficient is estimated in a separate OLS regression with patient and calendar-quarter fixed-effects. Robust standard errors clustered by patient in parentheses; *significant at 5%; **significant at 1%. Percentage changes are measured relative to baseline-year means. Dollar figures based on separate OLS regression where each dependent variable enters in levels. †Low self-control patients have a baseline compliance-index less than two.

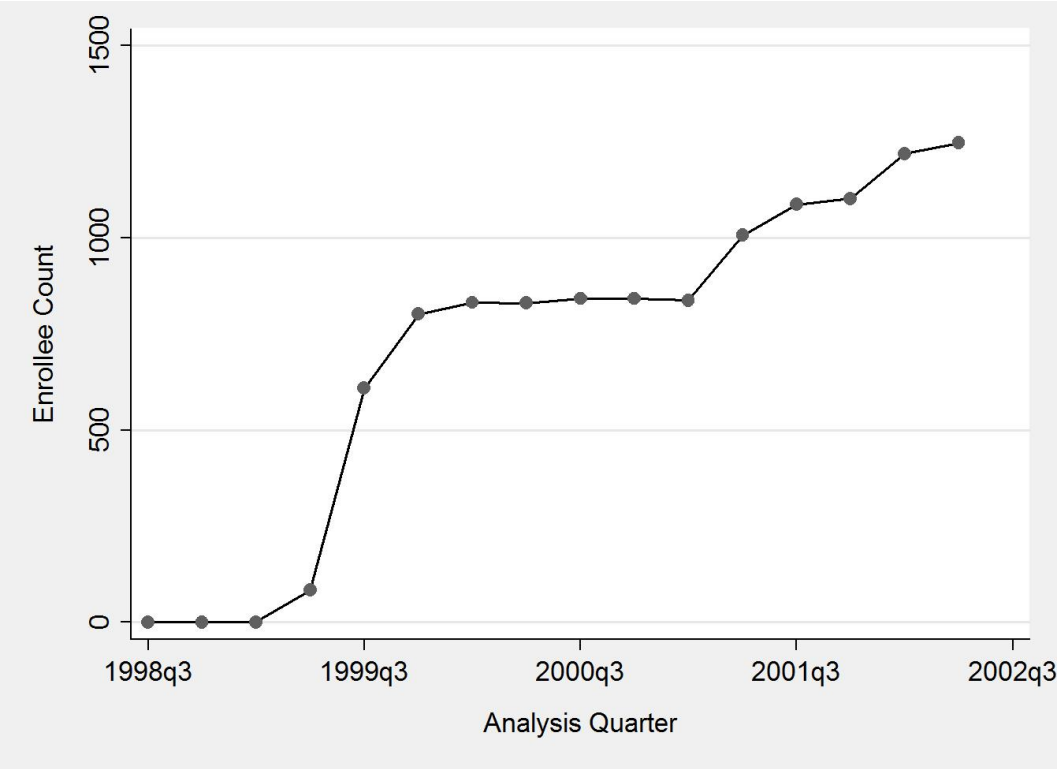

Figure 1: Enrollment in the Disease Management Program

Appendix A: Variable Definitions and Construction

Table A-1: Variable Definitions

	Baseline Characteristics	
Age	Fixed	Age in years at September 1999
Income	Fixed	Average household income from CPS for patient's Zip code
Medicare	Fixed [1/0]	Patient insured by Medicare [1/0]
Medicaid	Fixed [1/0]	Patient insured by Medicaid [1/0]
Baseline HbA1c	Fixed	Pre-intervention HbA1c Test results
High Self-Control	Fixed [1/0]	Compliance Index ≥ 2 during baseline year
Comorbidity	Fixed [1/0]	Diagnosed with cardiac condition (CAD, CHF, COPD) at baseline.
Attrition	Fixed [1/0]	Patient left Fallon prior to the end of teh study period.
	Compliance Measures	
HbA1c Compliant	Qtr [1/0]	Received HbA1c exam in last year
Eye Compliant	Qtr [1/0]	Received retinal exam in last year
Lipid Compliant	Qtr [1/0]	Received Lipids panel in last year
Malb Compliant	Qtr [1/0]	Received Kidney screening in last year
Compliance Index	Qtr	HbA1c Compliant + Eye Compliant + Lipid Compliant + Malb Compliant
	Health Outcomes	
HbA1c	Qtr	HbA1c test results
HbA1c ≥ 9.5	Qtr [1/0]	HbA1c results above 9.5
Inpatient visit	Qtr [1/0]	Inpatient admission
Morbidity	Qtr [1/0]	Diagnosed CAD, CHF, COPD, Stroke, Blindness, Renal failure, Amputation
Mortality	Qtr [1/0]	Died in current quarter
	Financial Outcomes	
Total Cost	Qtr	Dollar value of claims (Inpatient, outpatient, and pharmaceutical)
Inpatient Claims	Qtr	Dollar value of inpatient claims
Outpatient Claims	Qtr	Dollar value of outpatient claims
Pharma Claims	Qtr	Dollar value of pharmaceutical claims

Table A-2: Diagnostic Variable Coding

	Compliance Measures	
	ICD9 Codes*	CPT4 Codes†
HbA1c	Test results	83036
Kidney Exam	Test results OR 39.27, 39.42 39.43, 39.53, 39.93, 39.94, 39.95, 54.98, 250.4, 403, 404, 405.01, 405.91, 753.0, 753.1, 791.0	36800, 36810, 36815, 50300 ,50340, 50360, 50365, 50370, 50380, 90920, 90921, 90924, 90925, 90935, 90937, 90945, 90947, 90989, 90993, 90997, 90999
Lipids Exam	Test results	80061, 83716, 83721
Retinal exam	V72.0, 14.1-14.5, 95.04 95.11, 95.12, 95.16	67101, 67105, 67107-67110, 67112, 67141, 67145, 67208, 67218, 67227, 67228, 92002, 92004, 92012, 92014, 92018, 92019, 92225, 92226, 92230, 92235, 92240, 92250, 92260, 92287, 99204, 99205, 99214, 99215, 99242-99245
	Comorbidities	
	ICD9 Codes	CPT4 Codes
CAD	410-415	
CHF	398.91,402.01,402.11,402.91,404.01,404.11, 404.91,425.1,425.4,425.7,425.9,428.*,425	
COPD	491.2,492.*,493.2,496.*	
Amputation	84.1*	28810, 28820, 28825
Blindness	369.*	
Renal Failure	584.*	
Stroke	430, 431, 434, 436	

∗ ICD-9 codes are a standard set of diagnostic and procedural codes (primarily used for evaluation) developed and maintained by the National Center for Health Statistics (NCHS) and the Centers for Medicare and Medicaid Services (CMS).
† CPT-4 codes are a set of procedural codes (primarily used for billing) that is developed and maintained by the American Medical Association.

Appendix B: Supplemental Regressions

Table B-1: DM Impact by Enrollment Quarter (with Time Trend): 1998Q3 to 2000Q1

	Low Self-Control†			High Self-Control		
	Compliance Index	Hba1c >9.5	log Total Costs	Compliance Index	Hba1c >9.5	log Total Costs
Treated * Year	0.042 (0.021)*	-0.015 (0.023)	-0.021 (0.042)	-0.009 (0.015)	-0.011 (0.009)	-0.039 (0.020)*
Enrollment Quarter	0.020 (0.057)	0.084 (0.073)	-0.597 (0.184)**	0.053 (0.033)	-0.043 (0.029)	0.069 (0.060)
Quarter = +1	0.045 (0.090)	-0.082 (0.103)	-0.778 (0.233)**	0.122 (0.055)*	-0.085 (0.037)*	0.113 (0.084)
Quarter = +2	-0.137 (0.128)	-0.120 (0.130)	-0.954 (0.311)**	0.161 (0.076)*	-0.073 (0.047)	0.106 (0.108)
Quarter = +3	-0.234 (0.276)	-0.220 (0.220)	-1.298 (0.652)*	0.258 (0.129)*	-0.160 (0.070)*	0.110 (0.160)
Patient fixed-effects	Y	Y	Y	Y	Y	Y
Calendar-quarter FEs	Y	Y	Y	Y	Y	Y
Patient-quarters	13,410	2,414	13,410	25,133	10,275	25,133
Patients	1,979	962	1,979	3,653	2,944	3,653

Notes: OLS regression. Robust standard errors clustered by patient in parentheses; *significant at 5%; **significant at 1%. HbA1C models have fewer observations because tests do not occur quarterly. †Low self-control patients have a baseline compliance-index less than two.

Table B-2: Long-run Impacts for Balanced Panel (First wave and unenrolled patients)

	Compliance and Health Outcomes			
	Compliance Index	HbA1c ≥ 9.5	Inpatient Visit	Morbidity†
Low Self-control†† (1,190 Patients)	0.290 (0.062)**	-0.060 (0.045)	-0.009 (0.012)	-0.006 (0.005)
	57.22%	-20.66%	-15.51%	-22.51%
High Self-control (2,258 Patients)	0.033 (0.037)	-0.100 (0.018)**	-0.011 (0.008)	0.000 (0.006)
	1.78%	-51.75%	-17.84%	0.31%
	Financial Outcomes			
	log Total Claims	log Inpatient	log Outpatient	log Pharma
Low Self-control (1,190 Patients)	-0.604 (0.106)**	-0.074 (0.099)	-0.382 (0.135)**	-0.573 (0.110)**
	($185)	($65)	($83)	($36)
High Self-control (2,258 Patients)	-0.038 (0.045)	-0.089 (0.071)	0.024 (0.064)	-0.037 (0.061)
	($0)	($118)	+$119	($1)

Notes: Each coefficient is estimated in a separate OLS regression with patient and calendar-quarter fixed-effects. Robust standard errors clustered by patient in parentheses; *significant at 5%; **significant at 1%. Percentage changes are measured relative to baseline-year means. Dollar figures based on separate OLS regression where each dependent variable enters in levels. †Marginal effects and robust standard errors from a discrete-time logistic hazard model. ††Low self-control patients have a baseline compliance-index less than two.

Table B-3: DM Impact by Enrollment Quarter: 1998Q3 to 2001Q1 (Treated Patients Only)

	Low Self-Control†			High Self-Control		
	Compliance Index	Hba1c >9.5	log Total Costs	Compliance Index	Hba1c >9.5	log Total Costs
Quarter = -2	0.150 (0.043)**	-0.070 (0.037)	-0.043 (0.077)	0.008 (0.028)	-0.028 (0.018)	-0.075 (0.046)
Quarter = -1	0.090 (0.057)	-0.088 (0.037)*	-0.213 (0.086)*	-0.021 (0.036)	0.016 (0.022)	-0.088 (0.050)
Enrollment Quarter	0.069 (0.068)	-0.058 (0.042)	-0.437 (0.112)**	0.026 (0.041)	-0.045 (0.022)*	-0.091 (0.051)
Quarter = +1	0.080 (0.080)	-0.153 (0.041)**	-0.560 (0.132)**	0.070 (0.047)	-0.097 (0.023)**	-0.094 (0.056)
Quarter = +2	0.027 (0.091)	-0.199 (0.053)**	-0.786 (0.155)**	0.074 (0.051)	-0.107 (0.024)**	-0.141 (0.065)*
Quarter = +3	-0.061 (0.095)	-0.177 (0.058)**	-0.894 (0.158)**	0.061 (0.055)	-0.097 (0.028)**	-0.137 (0.073)
Quarter 4+	-0.041 (0.103)	-0.190 (0.063)**	-0.846 (0.166)**	-0.054 (0.065)	-0.146 (0.029)**	-0.243 (0.079)**
Patient fixed-effects	Y	Y	Y	Y	Y	Y
Calendar-quarter FEs	Y	Y	Y	Y	Y	Y
Patient-quarters	6865	2063	6865	15963	7158	15963
Patients	453	397	453	1061	967	1061

Notes: OLS regression. Robust standard errors clustered by patient in parentheses; *significant at 5%; **significant at 1%. HbA1C models have fewer observations because tests do not occur quarterly.
†Low self-control patients have a baseline compliance-index less than two.

Table B-4: Compliance Health and Financial Impacts: 1998Q3 to 2002Q2 (Treated Patients Only)

	Compliance and Health Outcomes			
	Compliance Index	HbA1c ≥ 9.5	Inpatient Visit	Morbidity†
Low Self-control (453 Patients)	-0.037	-0.065	-0.024	-0.011
	(0.054)	(0.027)*	(0.011)*	(0.007)
	-6.51%	-19.79%	-34.58%	-44.94%
High Self-control (1,061 Patients)	0.056	-0.076	-0.009	-0.009
	(0.032)	(0.016)**	(0.008)	(0.005)
	2.87%	-31.83%	-12.17%	-25.48%
	Financial Outcomes			
	log Total Claims	log Inpatient Claims	log Outpatient Claims	Log Pharma Claims
Low Self-control (453 Patients)	-0.535	-0.201	-0.469	-0.426
	(0.104)**	(0.094)*	(0.108)**	(0.105)**
	($453)	($200)	($229)	($24)
High Self-control (1,061 Patients)	-0.059	-0.075	-0.042	-0.014
	(0.040)	(0.073)	(0.054)	(0.046)
	($242)	($56)	($198)	+$12

Notes: OLS regression. Robust standard errors clustered by patient in parentheses; *significant at 5%; **significant at 1%. HbA1C models have fewer observations because tests do not occur quarterly. †Low self-control patients have a baseline compliance-index less than two.

Table B-5: Morbidity Logits (Marginal Effects)

	Low Self-Control		High Self-Control	
Sample	Wave 1 and Never Enrolled	All Eligible Patients	Wave 1 and Never Enrolled	All Eligible Patients
Pr[Diagnosed]	0.032	0.033	0.035	0.035
Treatment	-0.008 (0.005)	-0.009 (0.005)*	0.002 (0.003)	0.002 (0.003)
Age	0.001 (0.000)**	0.001 (0.000)**	0.001 (0.000)**	0.001 (0.000)**
Male	0.003 (0.003)	0.003 (0.003)	0.014 (0.003)**	0.015 (0.002)**
Urban	-0.014 (0.007)	-0.017 (0.007)*	-0.003 (0.006)	-0.004 (0.006)
White	-0.010 (0.028)	-0.011 (0.027)	0.022 (0.022)	0.023 (0.021)
log Income	-0.004 (0.009)	-0.007 (0.008)	-0.005 (0.007)	-0.008 (0.007)
Medicare	0.005 (0.006)	0.004 (0.005)	0.003 (0.004)	0.002 (0.004)
log Cost	-0.001 (0.001)	-0.002 (0.001)	0.003 (0.001)	0.003 (0.001)*
log Inpatient	-0.002 (0.001)	-0.002 (0.001)*	-0.004 (0.001)**	-0.004 (0.001)**
log Pharma	0.001 (0.001)	0.000 (0.001)	0.003 (0.001)*	0.002 (0.001)*
HbA1c Exam	-0.004 (0.002)*	-0.004 (0.002)*	-0.004 (0.001)**	-0.005 (0.001)**
Eye Exam	-0.000 (0.002)	0.000 (0.002)	0.001 (0.001)	0.002 (0.001)
Lipids Exam	0.005 (0.004)	0.003 (0.003)	0.004 (0.001)**	0.005 (0.001)**
Inpatient Admit	0.043 (0.012)**	0.051 (0.013)**	0.073 (0.010)**	0.077 (0.010)**
Patients	1,258	1,469	2,233	2,527
Patient-quarters	11,021	12,578	19,031	20,878

Notes: Discrete-time logistic hazard model. Table reports marginal effects and robust standard errors clustered by patient; *significant at 5%; **significant at 1%. All control variables except treatment measured at baseline year. Unreported calendar-quarter effects included in all models.

Table B-6: Mortality Logits (Marginal Effects)

	Low Self-Control		High Self-Control	
Sample	Wave 1 and Never Enrolled	All Eligible Patients	Wave 1 and Never Enrolled	All Eligible Patients
Pr[Death]	0.0073	0.0060	0.0079	0.0070
Treatment	-0.001	-0.000	-0.002	-0.001
	(0.002)	(0.001)	(0.001)	(0.001)
Age	0.000	0.000	0.000	0.000
	(0.000)**	(0.000)**	(0.000)**	(0.000)**
Male	0.002	0.002	0.004	0.003
	(0.001)*	(0.001)**	(0.001)**	(0.001)**
Urban	-0.001	-0.001	-0.000	0.000
	(0.002)	(0.002)	(0.002)	(0.002)
White	-0.002	-0.001	0.005	0.004
	(0.007)	(0.006)	(0.006)	(0.006)
log Income	0.002	0.001	0.000	0.000
	(0.002)	(0.002)	(0.002)	(0.002)
Medicare	0.003	0.002	0.002	0.002
	(0.002)	(0.002)	(0.001)	(0.001)
log Cost	0.001	0.001	0.002	0.002
	(0.000)	(0.000)	(0.000)**	(0.000)**
log Inpatient	-0.000	-0.000	-0.001	-0.001
	(0.000)	(0.000)	(0.000)**	(0.000)**
log Pharma	-0.000	-0.000	0.001	0.000
	(0.000)*	(0.000)*	(0.000)	(0.000)
HbA1c Exam	0.001	0.001	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Eye Exam	0.000	0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Lipids Exam	-0.001	-0.002	-0.002	-0.002
	(0.001)	(0.001)	(0.000)**	(0.000)**
CHF	0.008	0.007	0.009	0.008
	(0.002)**	(0.002)**	(0.002)**	(0.002)**
COPD	0.005	0.005	0.004	0.004
	(0.002)**	(0.002)**	(0.002)*	(0.001)*
CAD	-0.001	-0.001	-0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)
Hypertension	-0.002	-0.002	-0.001	-0.001
	(0.001)*	(0.001)*	(0.001)	(0.001)
Inpatient Admit	0.008	0.007	0.009	0.008
	(0.002)**	(0.002)**	(0.002)**	(0.002)**
Patients	1,714	1,979	3,252	3,653
Patient-quarters	17,598	20,124	33,603	37,067

Notes: Discrete-time logistic hazard model. Table reports marginal effects and robust standard errors clustered by patient; *significant at 5%; **significant at 1%. All control variables except treatment measured at baseline year. Unreported calendar-quarter effects included in all models.