WOLFF, CAROLYN M., Ph.D. Do Work Incentives Work? Three Essays on the Impacts for Physicians and Welfare Recipients. (2013)
Directed by Dr. David C. Ribar. 157 pp.

The three essays in this dissertation focus on the impacts of work incentives geared towards two very different segments of the labor market. The first essay, "Does Incentive Pay Alter Physician Effort? An Analysis of the Time and Treatment that Physicians Provide to Patients," examines the link between incentive pay and effort among a group of highly-skilled workers: physicians. The other two essays, "Exiting TANF in South Carolina after the Deficit Reduction Act" and "What Happened to Cash Assistance for Needy Families," focus on a group of generally low-skilled, low-wage workers: welfare recipients. "Exiting TANF in South Carolina after the Deficit Reduction Act" examines the impact of a recent welfare reform aimed at promoting employment and self-sufficiency on durations of welfare recipiency. "What Happened to Cash Assistance for Needy Families?" identifies trends in welfare recipiency and self-sufficiency over the past twenty years.

While a number of studies have attempted to measure the impact of financial incentives on physician behavior, none has examined the impact of performance-based incentive pay on broad measures of physician effort. In "Does Incentive Pay Alter Physician Effort? An Analysis of the Time and Treatment that Physicians Provide to Patients," I use newly available data from the National Ambulatory Medical Care Survey from 2006 through 2008 to estimate the effect of three specific types of performance-based incentive pay – productivity incentives, patient-centered incentives, and practice profiling incentives – on both the time physicians spend with patients and the intensity

with which physicians treat patients. Using a discrete factor approximation approach to control for the endogeneity of incentive pay, I am able to estimate the impact of these types of incentive pay on physician effort. I find that performance-based incentive pay is associated with physicians spending significantly less time with each patient. I also find some evidence that performance-based incentive pay impacts physicians' intensity of treatment.

The Deficit Reduction Act of 2005 (DRA) narrowed and standardized the work and work readiness activities that satisfy the work requirement of the Temporary Assistance for Needy Families (TANF) program. In "Exiting TANF in South Carolina after the Deficit Reduction Act," I use administrative data from South Carolina's TANF program and employ event history techniques with a difference-in-difference estimation framework to analyze the effect of this policy change. I find that the DRA's definition of work and work readiness activities reduced the likelihood of black recipients to exit the TANF program in South Carolina while increasing the likelihood of exit for non-black recipients. For blacks, this decrease in the hazard comes from a decrease in the likelihood of exit through employment. For non-blacks, the result stems from an increase in the hazards for administrative exits and for other income exits. I also find that the reform led to longer durations of TANF benefit receipt in South Carolina for black recipients and shorter durations of cash assistance for non-black recipients.

A primary goal of welfare reform since the early 1990's has been to increase the self-sufficiency of welfare recipients. The essay "What Happened to Cash Assistance for Needy Families?," coauthored with David. C. Ribar, examines trends in the

characteristics and outcomes for recipient families to determine if welfare recipients are becoming more self-sufficient. Using annual public use data on AFDC and TANF households from the Department of Health and Human Services, we find both positive and negative trends over the past twenty years. We find that the size of the caseload has decreased, the fraction of the caseload with earned income has increased, and the average earnings of welfare recipients has increased. On the other hand, we find that the fraction of child-only cases has increased, the caseload has disproportionately dropped the least-skilled households, average benefits fell faster than earnings grew, and the majority of households that exit TANF have no earnings.

DO WORK INCENTIVES WORK? THREE ESSAYS ON THE IMPACTS FOR PHYSICIANS AND WELFARE RECIPIENTS

by

Carolyn M. Wolff

A Dissertation Submitted to
the Faculty of The Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Greensboro
2013

Approved by

David C. Ribar
Committee Chair

APPROVAL PAGE

This dissertation written by Carolyn M. Wolff has been approved by the following committee of the Faculty of The Graduate School at The University of North Carolina at Greensboro.

Committee Chair David C. Ribar

Committee Members Charles Courtemanche

Stephen Holland

Christopher Swann

August 26, 2013
Date of Acceptance by Committee

August 19, 2013
Date of Final Oral Examination

ACKNOWLEDGMENTS

I would like to thank Charles Courtemanche, Stephen Holland, David Ribar, Christopher Swann, and seminar participants at the University of North Carolina at Greensboro, the Food and Drug Administration, and the University of South Carolina for valuable comments and suggestions. I would also like to thank Marilyn Edelhoch, Qiduan Liu, and Jim Maurer for providing detailed information on the TANF policies in South Carolina.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

The three essays in this dissertation focus on the impacts of work incentives geared towards two very different segments of the labor market. The first essay, "Does Incentive Pay Alter Physician Effort? An Analysis of the Time and Treatment that Physicians Provide to Patients," examines the link between incentive pay and effort among a group of highly-skilled workers: physicians. The other two essays, "Exiting TANF in South Carolina after the Deficit Reduction Act" and "What Happened to Cash Assistance for Needy Families," focus on a group of generally low-skilled, low-wage workers: welfare recipients. "Exiting TANF in South Carolina after the Deficit Reduction Act" examines the impact of a recent welfare reform aimed at promoting employment and self-sufficiency on durations of welfare recipiency. "What Happened to Cash Assistance for Needy Families?" identifies trends in welfare recipiency and self-sufficiency of welfare recipients over the past twenty years.

While a number of studies have attempted to measure the impact of financial incentives on physician behavior, none has examined the impact of performance-based incentive pay on broad measures of physician effort. In "Does Incentive Pay Alter Physician Effort? An Analysis of the Time and Treatment that Physicians Provide to Patients," I use newly available data from the National Ambulatory Medical Care Survey from 2006 through 2008 to estimate the effect of three specific types of performance-

based incentive pay – productivity incentive pay, patient-centered incentive pay, and practice profiling incentive pay – on both the time physicians spend with patients and the intensity with which physicians treat patients. Each of these measures of performance-based incentive pay rewards a physician with a payment for meeting the specified performance goal.

Using a discrete factor approximation approach to control for the endogeneity of incentive pay, I am able to estimate the impact of these types of incentive pay on physician effort. I find productivity incentive pay is associated with physicians spending approximately two minutes (9.4 percent) less with each patient, on average. If a physician reduces his time spent with each patient from the average of 21 minutes and 47 seconds to 19 minutes and 43 seconds, in an eight hour workday he will see two additional patients. If, however, time spent with the patient is a proxy for quality, productivity incentive pay that increases overall physician effort by decreasing per-patient effort comes at the expense of the quality of medical care received by the patient. I find little evidence that patient-centered or practice profiling incentive pay impacts the amount of time a physician spends with his patients, but I do find evidence of unobserved heterogeneity leading to positive selection bias in the coefficient estimates for patient-centered and practice profiling incentive pay in the model for time spent with patients.

I also find some evidence that performance-based incentive pay impacts physicians' intensity of treatment. Patient-centered incentive pay appears to reduce treatment intensity while practice profiling incentive pay seems to increase treatment intensity.

The Deficit Reduction Act of 2005 (DRA) narrowed and standardized the work and work readiness activities that satisfy the work requirement of the Temporary Assistance for Needy Families (TANF) program. In "Exiting TANF in South Carolina after the Deficit Reduction Act," I use administrative data from South Carolina's TANF program and employ event history techniques with a difference-in-difference estimation framework to analyze the effect of this policy change. In particular, I examine the impact of the reform on the likelihood of exit from the TANF program and on the paths to exit from the TANF program in South Carolina. Using these results, I conduct simulations to determine how the reform impacted the duration of benefit receipt in South Carolina.

I find that, during the first four months of a spell, the DRA's definition of work and work readiness activities reduced the likelihood of black recipients to exit the TANF program in South Carolina while increasing the likelihood of exit for non-black recipients. I also find that the reform reduced the likelihood of exit due to employment for blacks at the beginning of a spell while increasing the likelihood of exit due to other sources of income or due to administrative reasons for non-blacks. Finally, I find that the reform led to longer durations of TANF benefit receipt in South Carolina for black recipients and shorter durations of cash assistance for non-black recipients.

The primary goal of welfare reform since the early 1990's has been to increase the self-sufficiency of welfare recipients. The essay "What Happened to Cash Assistance for Needy Families?" examines trends in the characteristics and outcomes for recipient families to determine if welfare recipients are becoming more self-sufficient. Using annual public use data on AFDC and TANF households from the Department of Health

and Human Services, we find both positive and negative trends over the past twenty years. We find that the size of the caseload has decreased, the fraction of the caseload with earned income has increased, and the average earnings of welfare recipients has increased. On the other hand, we find that the fraction of child-only cases has increased, the caseload has disproportionately dropped the least-skilled households, average benefits fell faster than earnings grew, and the majority of households that exit TANF have no earnings.

We also find that the general well-being of those at risk of becoming dependent on welfare increased during the 1990s but declined from 2000 to 2008. The poverty rate of single mothers declined from 47 percent in 1991 to 33 percent in 2000 but increased to 37 percent by 2008. The employment rate of single mothers increased from 56 percent in 1990 to 63 percent by 200 but dropped to about 60 percent by 2008. Finally, of the 23.7 million people living below the poverty line in 2008, TANF only reached 3.7 million.

CHAPTER II

DOES INCENTIVE PAY ALTER PHYSICIAN EFFORT? AN ANALYSIS OF THE TIME AND TREATMENT THAT PHYSICIANS PROVIDE TO PATIENTS

1. Introduction

With the recent trend towards pay-for-performance initiatives in healthcare, the American Medical Association has pushed for physicians to take a more active role in ensuring they receive the appropriate share of pay-for-performance bonuses (Elliott, 2012). As physicians alter their effort in response to these incentives, it is likely to impact the cost and quality of medical care. This paper examines the effect of performance-based incentive pay on quantitative measures of physician effort: the amount of time that a physician spends with a patient and two measures of the intensity with which the physician treats the patient. A decrease in either of these components of per-patient physician effort would be associated with a decrease in the cost of medical care and, potentially, a decrease in the quality of medical care.

A large breadth of literature in both the economics and healthcare fields attempts to measure the impact of financial incentives on physician behavior, and much of this literature falls naturally into two groups.[1] The first group of studies examines the impact of physicians' quality of care (QOC) and productivity performance-based incentive pay (i.e. rewarding physicians with higher pay for meeting some performance goal) on

[1] For a review of studies on the impact of incentive pay in other labor markets, see Prendergast (1999) and Lazear (2000).

process-of-care outcomes, such as rates of immunization or screening for disease; however, in focusing on very specific process of care outcomes, these studies can only account for the impact of incentive pay on a limited component of physician effort. A second group of studies examines the effect of reimbursement methods (i.e. the methods by which insurers pay for services rendered) on broader measures of physician effort; however, reimbursement incentives are only one of many types of financial incentives that physicians face.[2]

What is lacking in the literature is a study of the impact of performance-based incentive pay on broad measures of per-patient physician effort. We know from the literature on reimbursement incentives that physicians do respond broadly to financial incentives; however, many physicians work as salaried employees in medical groups. These salaried group physicians are shielded from reimbursement incentives, but they are subject to the group's performance-based incentive pay. Further, even incentive pay targeted at narrowly defined aspects of physician effort may have spillover effects, resulting in a broader impact on physician effort than intended. Thus, it is important to consider the broad impacts of incentive pay on physician effort.

In this paper, I examine the effect of three types of performance-based incentive pay on the amount of time physicians spend with each patient and on two measures of the

[2] Three common forms of reimbursement are capitation, fee-for-service (FFS), and salary. Under capitated reimbursement, the physician or medical group receives a set monthly payment for each patient on the roster, regardless of the services provided. Under FFS reimbursement, the physician or medical group is paid a previously agreed upon fee for each service rendered. Under salary reimbursement, the physician receives a fixed salary regardless of the number of patients under his care or the services rendered for those patients.

intensity of the physicians' treatment. Examining both the time and intensity of treatment is critical to understanding the full effect of performance-based incentive pay. Further, I model the selection of physicians into performance-based incentive pay using a discrete factor approximation approach, thus addressing a source of selection bias that may be present due to the endogeneity of incentive pay. I find strong evidence that performance-based incentive pay reduces the amount of time physicians spend with each patient. I also find weaker evidence that incentive pay impacts physicians' intensity of treatment.

This paper contributes to the literature in three ways. First, I examine the impact of performance-based incentive pay, including two types of incentive pay not considered by previous studies, on three broad measures of physician effort. Second, I use newly available data from the National Ambulatory Medical Care Survey (NAMCS) on the performance-based incentive pay received by each physician as well as data on the effort exerted by the physicians in treating each patient. Third, I employ a discrete factor methodology to control for unobserved heterogeneity that would otherwise bias the results.

2. Literature

In their seminal paper, Gaynor and Pauly (1990) find that when the medical group compensation structure provides physicians with the incentive to see more patients, physicians respond by increasing the number of patients they treat each week. Many subsequent empirical studies have attempted to measure the impact of different types of

financial incentives on physician effort.[3] One body of literature examines the impact of physicians' performance-based incentive pay on process-of-care outcomes, with a focus on two types of performance-based incentive pay: that targeted at improving quality of care outcomes and that targeted at improving productivity. A second body of literature examines the impacts of reimbursement incentives on broader measures of physician effort, such as time spent with patients and intensity of treatment. I review each of these in turn.

Quality of care performance-based incentive pay consists of financial rewards paid to physicians who meet some clinical performance or quality goal, often associated with preventive care (e.g. rates of immunization or disease screening). The literature on physician QOC performance-based incentive pay has generated equivocal evidence that this type of incentive alters physician effort.[4] For instance, Fairbrother et al. (1999, 2001) find that QOC incentive pay rewarding a physician for childhood immunization leads to better documentation of immunizations received outside the physician's practice but has little impact on the number of immunizations actually administered by the physician. Beaulieu and Horrigan (2005) find that QOC incentive pay combined with education and healthcare management tools is associated with greater levels of physician effort for diabetes patients. Finally, Grady et al. (1997) find no evidence that physician QOC incentive pay impacts mammography referral rates.

[3] In addition, Grumbach et al. (1998) show that financial incentives have a psychological impact on physicians.

[4] A larger body of literature on the impact of QOC performance-based incentive pay for medical groups on individual physician effort has also found weak results. See Petersen et al. (2006) or Rosenthal and Frank (2006) for a review as well as Campbell et al. (2007), Mullen et al. (2010), and Rosenthal et al. (2005).

Productivity performance-based incentive pay rewards physicians for providing a greater quantity of services (e.g. office visits). Wee et al. (2001) find that productivity performance-based incentive pay is associated with physicians providing lower levels of some types of preventive care (such as Pap smears and cholesterol screens) but has no impact on other processes of preventive care. A common weakness of Beaulieu and Horrigan (2005) and Wee et al. (2001) is that these studies may suffer from multiple sources of bias, including unobserved heterogeneity, and thus caution should be taken in interpreting the results as causal.

The studies of reimbursement incentives provide stronger evidence that physicians alter their effort in response to these incentives, as reimbursement schemes with fewer positive incentives, such as capitation, are found to be associated with lower amounts of overall and per-patient physician effort. For example, both Balkrishnan et al. (2002) and Melichar (2009), using patient-level data from the NAMCS, find that physicians spend about one minute less with capitated patients than with other patients.[5] Similarly, many studies find that fee-for-service (FFS) physicians conduct more patient visits than those under a reimbursement system with fewer positive incentives (Devlin and Sarma, 2008; Grytten et al., 2009; Sarma et al., 2010; and Sorensen and Grytten, 2003). Shafrin (2010) finds that reimbursing a specialist via FFS is associated with a greater number of surgeries, and Dumont et al. (2008) find that when physicians switch from FFS to a less incentive-based reimbursement scheme, they reduce their volume of

[5] It is interesting that while Melichar (2009) uses fixed effects to control for physician selection into reimbursement schemes and Balkrishnan et al. (2002) do not account for such selection, both studies find the same effect.

clinical services provided. On the other hand, Balkrishnan et al. (2002) find that physicians provide more counseling and education services for capitated patients. In a related study, Glied and Zivin (2002) find that the average time spent with a patient is shorter in HMO-dominated practices; however, the number of tests and number of medications provided is larger.

It is clear from the literature that physicians respond broadly to reimbursement incentives by altering their overall and per-patient levels of effort; however, it is still unclear as to whether and how physicians respond to other types of financial incentives, specifically performance-based incentive pay. There is a lack of consensus in the literature on the effect of QOC and productivity performance-based incentive pay on quantitative process of care outcomes. Further, to my knowledge, there has been no previous examination of the effect of QOC and productivity performance-based incentive pay on broad measures of per-patient physician effort nor has there been much study of the effects of practice profiling or patient satisfaction performance-based incentive pay on any form of physician effort. A clear understanding of these effects, as I provide in this study, may aid in interpreting the impact of physicians' incentive pay on the cost and quality of healthcare in the United States.

3. Conceptual Framework

There are a number of reasons why a physician might face performance-based incentive pay, and each of these has a unique implication for the effect of performance-based incentive pay on physician effort. In this section, I develop a conceptual framework demonstrating why a medical group may want to contract with physicians on

effort. I then propose three applications of performance-based incentive pay in this framework and discuss the implications for physician effort in each application.

Consider an individual physician employed by a medical group, and suppose that the physician's objective is to maximize his utility, which is increasing in income and decreasing in effort. Also suppose that the medical group's objective is to maximize profit, which is a function of the medical services provided by each physician and the income paid to each physician employed by the group. The medical services provided by a physician may include office visits as well as any tests or procedures performed by the physician. The quantity of medical services provided is an increasing function of the physician's overall level of effort; thus the medical group's profit is increasing in overall physician effort.[6]

Suppose that overall physician effort may be divided into two components: the number of patients treated and the intensity of the treatment. Each of these components will have a separate impact on the medical group's profit. Consider first the number of patients treated, assuming that hours worked each day is exogenous, that the price to patients of an office visit is independent of physician effort, and that the physician is operating under a full schedule (i.e. the physician does not have any extra time in his

[6] The distinction between "overall" and "per-patient" physician effort is important here. Gaynor and Pauly (1990) assume that a physician can increase the medical services provided, and thus increase revenue, by increasing his per-patient level of effort. While this may be true for solo practitioners, physicians working in a medical group often have exogenous constraints placed on them by the group, and these constraints may limit a physician's ability to raise revenue through increases in per-patient effort. This is particularly relevant when thinking about effort in terms of the physician's time spent with patients.

workday to see additional patients without decreasing the time spent with each patient).[7] In this case, the physician may be able to increase his overall level of effort by decreasing his per-patient level of effort: reducing the amount of time spent with each patient so as to see more patients each day.[8] By seeing more patients each day, the physician has increased the quantity of medical services provided, leading to an increase in the group's profit.

Now consider the second component of physician effort: the intensity of treatment. In this paper, I think of intensity of treatment as how aggressive the physician is in treating the patient. Does the physician order a battery of tests or perform any procedures? Many of the services embodied in treatment intensity carry a price in addition to the price for an office visit; thus, increasing the intensity of treatment will increase the medical services provided and consequently, the group's profit.

Given the above assumptions, a profit-maximizing medical group will have a keen interest in the effort levels exerted by the physicians it employs. As such, there are a number of reasons as to why a medical group may use incentive pay to try to alter the physicians' levels of effort, and I discuss three of these reasons in the following

---

[7] Hours of work can be thought of as being exogenous if one assumes that it is decided on by the group rather than by the individual physician. This assumption seems reasonable as the group, not the individual physician, is choosing the levels of non-physician labor and capital, which the individual physician takes as exogenous. Just as the group may choose to employ the receptionist for eight hours each day, the group may choose to keep the medical office suite open for eight hours each day, which limits the time a physician can work to only eight hours. This assumption is further supported by the inability of Gaynor and Pauly (1990) to reject the exogeneity of physician hours in the production function.

[8] This result relies on the assumption that the physician is operating under a full schedule. If, on the other hand, the physician has idle time during his workday in which he is not seeing patients, he could increase the number of patients treated without decreasing his time spent with each patient.

subsections. This list is not meant to be exhaustive but simply to illustrate the potential uses of incentive pay in medical group practices.

3.1 Moral Hazard

The first reason for why a medical group may use incentive pay is to correct for moral hazard. For example, if the group could contract with the physician on income and effort, the optimal level of physician effort could be reached; however, physician effort may be subject to discretion on the part of the physician based on patient characteristics, a patient's history with the physician, and physician preferences, making contracting on effort difficult in the healthcare field. This difficulty could present the physician with the opportunity to shirk, providing a level of effort that is not optimal from the medical group's perspective. It may be possible for a medical group to use incentive pay to reduce this moral hazard and better align the objectives of the group and the physician so that the level of effort exerted by the physician is closer to the optimal level. The practice could pay the physician a performance-based incentive for certain verifiable types of revenue-generating effort, such as the number of patients seen each day (e.g. a productivity incentive). This incentive will offset some of the disutility associated with the increase in overall effort while inducing the physician to spend less time with each patient so as to see more patients each day.

3.2 Patient Well-being and Induced Demand

A second reason for why a medical group may use incentive pay is to induce demand. Suppose the group's objective is as before but now the physician cares about the well-being of his patients. In this case, the physician's objective is to maximize

utility with respect to income and patient well-being, where his utility is increasing in both income and patient well-being. The medical group, on the other hand, is only concerned about profit and prefers that patients consume greater quantities of medical services without consideration for patient well-being. If providing medical care to optimize the well-being of the patient is not consistent with profit maximization, the medical group may use incentive pay tied to the number or pattern of profitable medical services that a physician provides to his patients (e.g. a productivity or practice profiling incentive), inducing the physician to provide more revenue-generating medical services, such as office visits (by reducing the amount of time spent with each patient), diagnostic and screening services, and procedures.

3.3 Patient Satisfaction

A third reason for why a medical group may use incentive pay is to increase patient satisfaction. In seeking to maximize the current and expected future stream of profit, suppose the medical group cares about patient satisfaction, as more satisfied patients will be more likely to return to the practice and consume additional medical services in the future. Also suppose that the physician does not care about patient satisfaction.[9] If patient satisfaction is an increasing function of per-patient physician effort, the objectives of the practice and the physician will be at odds. The medical group could provide the physician with incentive pay tied to current patient satisfaction (e.g. a patient satisfaction or quality of care incentive) in order to offset his disutility associated

[9] This does not preclude the physician from wanting to maximize the patient's well-being. Rather it assumes that what is in the patient's best interest medically and what the patient prefers are not necessarily the same.

with effort. This could induce the physician to provide a higher level of per-patient effort by spending more time with each patient, providing more diagnostic and screening services, and performing more procedures.

3.4 Implications

Each of these motivations for incentive pay has implications for physician effort. First, productivity incentive pay aimed at resolving moral hazard will serve to offset some of the disutility associated with an increase in overall physician effort. Thus, productivity incentive pay should reduce the amount of time a physician spends with each patient so that the physician is able to see more patients each day. Second, productivity or practice profiling incentive pay aimed at promoting induced demand will induce the physician to provide more revenue generating services. This should also reduce the amount of time a physician spends with each patient so as to increase the number of office visits, and it should increase the intensity of treatment so as to increase the quantity of other services provided (such as diagnostic and screening services and procedures). Finally, patient satisfaction or quality of care incentive pay aimed at increasing patient satisfaction should cause the physician to increase his effort in ways that patients value. If patients value time with the physician and greater treatment intensity, the physician will respond by increasing these components of his effort.

4. Data

The data used in this essay come from the Centers for Disease Control and Prevention's 2006, 2007, and 2008 National Ambulatory Medical Care Surveys (NAMCS), which randomly samples and interviews office-based physicians to collect

information on the physicians and on a sample of the patients they treat during a one-week reporting period.[10] The data collected on physicians include information such as employment status, the types of incentive pay received, and the services provided by the medical group. The information collected on the physicians' patients comes directly from the patients' charts and includes the number of minutes the physicians spent with each patient and the services ordered or performed by the physicians as well as information on the patient's health and demographics. These data are particularly well suited for this analysis because they not only contain information on per-patient physician effort from the patients' charts, but they also contain data on the types of incentive pay physicians receive, available only in the 2006, 2007, and 2008 survey years. Combining the three survey years of NAMCS data yields an initial dataset with 3,811 physicians and 90,911 patient-level observations.

A challenge to analyzing the data in this analysis, however, is that they are observed at two different levels: the measures of physician effort are observed at the patient-level while the incentive pay variables are observed at the physician-level. Because I do not observe patient-specific incentive pay, I aggregate the patient-level data up to the physician level by averaging over all patients for each physician. If the physician receives incentive pay for meeting a targeted level of effort across (or averaged over) all patients, then this approach is appropriate because one would not expect the

[10] The sampling procedure uses the physician's expectations over the number of days he will see patients and the total number of patients he will see during the week to select approximately 30 patient visits from each sampled physician. However, due to error in the physicians' expectations and non-participation, approximately 24 patient record forms were completed on average for each physician from 2006 to 2008.

response to vary systematically across patients. On the other hand, if this assumption is incorrect and physicians do indeed face different incentive pay for different groups of patients (e.g. patients with particular insurance plans), this model assumes a uniform response to incentives across all patients treated rather than a patient-type-specific response.[11] One example where assuming a uniform response may not be appropriate is incentive pay that differs across patients with public versus private insurance due to the differences that exist in reimbursement; thus, I model the impact of performance-based incentive pay on physician effort separately for private insurance, Medicare, and Medicaid patients.

I exclude from the analysis to patients who pay for their care out-of-pocket, as my assumption regarding the patient's price of an office visit being independent of physician effort may not be valid for patients who pay for their care out-of-pocket. I omit patients from the sample if they saw a provider other than the physician, as the physician's incentive pay can have no effect on his effort if he is not the one treating the patient. The 2,647 patients in the sample who saw a provider other than the physician have a time value of zero; dropping these observations removes the left-censoring of this outcome variable. I also omit the 23,395 patient-level observations where time spent with the

[11] This is more likely to be the case if the incentive payments are paid by a third-party payer rather than by the medical group. This may not be cause for concern, however, as Glied and Zivin (2002) find that physicians tend to respond to the financial incentives associated with the typical patient rather than to those associated with each individual patient. Further, for patient-type-specific incentive payments to impact only the targeted patients, the physician must be able to identify the patients' type. Given that the medical office staff, not the physician, is usually responsible for billing, it is unlikely that the physician knows a patient's type before he treats the patient. Knowing that these patient-type-specific incentive payments may have a limited targeted impact in this case, they are likely to be less widely utilized by medical groups.

patient is imputed to limit the amount of measurement error in the dependent variable and to reduce the incidence of artificial associations. I further limit my sample to the 2,687 physicians working in group practices as my assumption regarding the exogeneity of hours worked will not be valid for solo practitioners. Finally, I drop any observations with missing values, resulting in a final sample of 1,930 physicians and 37,404 patients.

4.1 Physician Effort Measures

The NAMCS collects information on three measures of per-patient physician effort that I use in the analysis. The first measure of effort is the number of minutes that the physician spent with the patient which, after aggregation, results in a variable for the sample average number of minutes spent with each patient. Due to heaping that occurs when the physician records how much time he spent with the patient, there are spikes in the distribution of this aggregated variable at regular intervals. On average, physicians in the sample spend just under 22 minutes with each patient.

The two remaining measures of per-patient physician effort are indicators of treatment intensity. The first is the count of diagnostic and screening services provided, which captures the number of examinations performed as well as any imaging, blood tests, scope procedures, or other tests ordered or provided for the patient. The second measure of treatment intensity is the number of procedures performed on the patient. Aggregating these measures up to the physician level results in variables for the sample average number of diagnostic and screening services provided and the sample average number of procedures performed. As with the aggregated time variable, the distributions

of these variables exhibit heaping at regular intervals. On average, physicians provide 3.6 diagnostic and screening services and perform 0.2 procedures for each patient.

4.2 Incentive Pay Measures

The 2006, 2007, and 2008 waves of the NAMCS asked each physician if any of four types of performance-based incentive pay were taken into account in the physician's base pay, bonuses, or withholds. First, productivity incentive pay rewards physicians based on productivity measures, such as the number of cases seen each day. The second type of incentive pay that physicians are asked about is a reward based on quality of care, where quality of care could include, for example, the number of preventive services performed. The third type of incentive pay rewards physicians based patient satisfaction as indicated in patient surveys. Fourth, practice profiling incentive pay rewards physicians based on the patterns of services, such as lab tests, procedures, and referrals, provided.

Quality of care and patient satisfaction incentive pay are both patient-centered in that they reward physicians based, to some extent, on patient outcomes. The QOC measure rewards physicians for providing preventive care, which reduces the patient's risk of experiencing an adverse health shock. The patient satisfaction measure rewards physicians for providing patients with the services they value. If patients value preventive care and the associated reduction in risk, patient satisfaction incentive pay will reward physicians for those aspects of medical care that are also rewarded by QOC incentive pay. Recognizing this, medical groups are likely to utilize both patient satisfaction and QOC measures, possibly in a single incentive pay instrument. Indeed, I

find that of the 338 physicians in the sample who receive patient satisfaction incentive pay, only forty-three (13%) do not also face QOC incentive pay. Further, of the 394 physicians in the sample who face QOC incentive pay, only ninety-nine (25%) do not also face patient satisfaction incentive pay. Due to this low incidence, a multivariate analysis cannot precisely distinguish the individual effects of both types of incentive pay simultaneously. To account for this, I combine these incentive pay variables into a single binary indicator variable for either type of patient-centered incentive pay.[12]

4.3 Exogenous Explanatory Variables

The NAMCS data contain a rich set of variables that I employ as controls in the analysis. I use information on the patient's age, race, ethnicity, gender, and primary method of payment to control for patient characteristics. I use data on the patient's reason for the office visit and whether the patient has been previously seen in the medical practice to control for characteristics related to the patient's health. To control for physician characteristics, I include data on whether the physician is the patient's primary care provider, whether the physician is a medical or osteopathic doctor, whether the physician is an owner in the group, whether the physician is accepting new patients, whether the physician sees patients on evenings or weekends, and information of the physician's specialty. I also use information on whether the medical practice performs its own lab testing, whether the practice uses electronic medical records, and the type of medical practice to control for medical group characteristics. I use data on geographic

[12] Sensitivity analyses show that the results are not sensitive to this restriction.

region and whether or not the practice is located in an MSA to control for regional effects. Table A1 of the Appendix presents sample statistics for these variables.

4.4 Descriptive Statistics

Table 1 shows the average per-patient physician effort outcomes grouped by the incentive pay variables. Overall, these descriptive statistics suggest that physicians who face incentive pay exert different levels of effort than other physicians and that the type of incentive pay matters. Physicians who receive productivity incentive pay spend less time, on average, with their patients, supporting the idea that productivity incentive pay may be used to correct for moral hazard. Physicians who face patient-centered or practice profiling incentive pay provide significantly more diagnostic and screening services, on average, which is consistent with the use of incentive pay to induce demand.

Physicians who face patient-centered incentive pay perform fewer procedures, on average. This is contrary to induced demand, although one may expect these results if this type of incentive pay is targeted towards a group of physicians, such as primary care physicians, who typically perform fewer procedures than their colleagues in other specialties. Indeed, I find that primary care physicians are more likely to not perform procedures during the reporting period and are more likely to receive patient-centered incentive pay than physicians who practice in medical or surgical specialties.

It is important to keep in mind that there are likely observable and unobservable characteristics confounding these results. More advanced multivariate techniques must be employed before a definitive statement can be made about the causal effects of incentive pay on per-patient physician effort.

5. Econometric Methods

In this study, I employ multivariate techniques to estimate the relationship between the incentive pay faced by physicians and three quantitative measures of average per-patient physician effort: (1) the number of minutes spent with each patient; (2) the number of diagnostic and screening services provided to each patient; and (3) the number of procedures performed on each patient.

For the baseline model, I estimate the following equation using Ordinary Least Squares (OLS) for the non-left-censored outcome (minutes spent with each patient) and a Tobit model for the left-censored outcomes (number of diagnostic and screening services and number of procedures):

$$y_0 = \beta_0 + \beta_1 ProdIncPay + \beta_2 PatIncPay + \beta_3 ProfIncPay + \beta_4 X + \varepsilon_0 \quad (1)$$

where y_0 is a quantitative measure of average per-patient physician effort, *ProdIncPay* is a binary variable for productivity incentive pay, *PatIncPay* is a binary variable for patient-centered incentive pay, *ProfIncPay* is a binary variable for practice profiling incentive pay, and X is the full set of control variables discussed above, expressed as a single vector for simplicity. A Tobit model is necessary when the outcome is the number of diagnostic and screening services or the number of procedures because during the reporting period, thirty-six physicians (2%) do not provide any diagnostic and screening services and 770 physicians (40%) do not perform any procedures.

While the baseline model controls for the effect of observed confounders on per-patient physician effort, it fails to account for the effect of unobserved confounders.

Suppose that the error term in equation (1) is generated as

$$\varepsilon_0 = \rho_0 v + u_0 \tag{2}$$

where u_0 is normally distributed, v is an unobserved factor, and ρ_0 is a factor loading term. If the unobservable physician characteristic, v, influences both the physician's selection into a medical group with incentive pay and the effort exerted by the physician, the OLS and Tobit estimators will be biased. I control for this selection bias by modeling the relevant unobserved heterogeneity through discrete factor approximation which will represent this unobservable physician characteristic with a discrete distribution.[13]

Mroz (1999) presents the discrete factor approximation framework for the case of a single dummy endogenous variable. Following his framework, suppose that whether or not the physician chooses each of the three types of incentive pay is determined by an underlying latent variable according to

$$y_1^* = \alpha_{10} + \alpha_{11}X + \rho_1 v + u_1 \tag{3}$$

$$ProdIncPay = \begin{cases} 1 & if\ \ y_1^* \geq 0 \\ 0 & otherwise. \end{cases}$$

$$y_2^* = \alpha_{20} + \alpha_{21}X + \rho_2 v + u_2 \tag{4}$$

[13] Since the NAMCS data are collected as a repeated cross section of physicians, a fixed effects estimator as used by Melichar (2009) cannot be employed here.

$$PatIncPay = \begin{cases} 1 & if\ \ y_2{*} \geq 0 \\ 0 & otherwise. \end{cases}$$

$$y_3{*} = \alpha_{30} + \alpha_{31}X + \rho_3 v + u_3 \tag{5}$$

$$ProfIncPay = \begin{cases} 1 & if\ \ y_3{*} \geq 0 \\ 0 & otherwise. \end{cases}$$

As in equation (2), u_1, u_2, and u_3 are normally distributed; v is an unobserved factor approximated by a discrete distribution with two points of support; and ρ_1, ρ_2, and ρ_3 are factor loading coefficients. [14] Only the outcomes y_0, *ProdIncPay*, *PatIncPay*, and *ProfIncPay* and the controls in X are observed. The parameters β_0, β_1, β_2, β_3, β_4, α_{10}, α_{11}, α_{20}, α_{21}, α_{30}, α_{31}, and ρ_0 from equations (1) – (5), along with one point of support and weight from the discrete distribution of v, are estimated simultaneously.[15]

As mentioned previously, the dependent variables exhibit heaping at regular intervals. To correct for this in the discrete factor approximation models, I discretize the dependent variables into ordered variables with equal-sized bins according to the heaping

[14] I tested the models with additional points of support. The models converge with three points of support when ρ_2 and ρ_3 are both fixed at one; however, the models fail to converge with three points of support when when ρ_2 and ρ_3 are estimated freely or when the sample is divided to test for heterogeneity. In the models with three points of support and ρ_2 and ρ_3 both fixed at one, the probability on the third point of support is less than 0.04 in each model, suggesting that the third point of support carries little weight. Further, the conclusions do not change when the third point of support is included. Therefore, for consistency, I present results of all models estimated with two points of support.

[15] I estimate these equations jointly using the aML software (Lillard and Panis, 2003).

and estimate equation (1) as an ordered probit model with known thresholds.[16, 17] With a large number of thresholds, this approach approximates the estimates from a continuous model; and constructing the dependent variables as ordered variables with equal-sized bins according to the spikes in the distributions corrects for the non-normality of these variables.

The principle advantage of the discrete factor approximation approach is that it allows for a flexible specification of the unobserved heterogeneity. As Heckman and Singer (1984) show, if one were to choose a specific distribution for the unobserved factor, such as a normally distributed random effect, the parameter estimates of the model would be sensitive to that choice. Discrete factor approximation allows one to avoid such sensitivity of the parameter estimates. However, this approach is not without its limitations. Primarily, discrete factor approximation assumes that there is a single, time-invariant, unobserved variable that is the source of the omitted variable bias. In this case, that single unobservable could be characterized as a measure of physician work habits.

A further limitation of the model is that identification of the full set of parameters relies on the nonlinear specification of the model, on the restriction of $\rho_1 = 1$ in the variance-covariance matrix, and on the independence of v and X. While this identification is achieved in theory, the model may not be identified in practice. Indeed, this is the case for the censored outcomes. Thus, I present results from the model with ρ_2

[16] In discretizing the left-censored outcomes, I create a unique bin for the zero values to maintain the Tobit structure.

[17] Estimating equation (1) as an ordered probit with known thresholds also allows me to circumnavigate an apparent bug in aML when estimating multi-equation models with both discrete and continuous outcomes.

and ρ_3 both fixed at one. This restriction allows identification of the model at the expense of restricting the unobserved variable to have the same correlation with each type of incentive pay.

6. Results

6.1 OLS and Tobit Models

Table 2 presents the estimates of equation (1) without controls for unobserved heterogeneity. As column (1) shows, productivity incentive pay has a strong impact on the time spent with patients. Physicians who face productivity incentive pay spend, on average, a statistically significant one minute and 44 seconds less with each patient than physicians who do not face this type of incentive pay. The estimates also suggest that patient-centered and practice profiling incentive pay are each associated with approximately one more minute spent with each patient, although these two estimates are not statistically significant.

Columns (2) and (3) present evidence that incentive pay also impacts treatment intensity. Physicians who receive practice profiling incentive pay order an average of 0.26 more diagnostic and screening services and perform 0.04 more procedures. These results suggest that physicians who face practice profiling incentive pay provide seven percent more diagnostic and screening services and perform twenty-two percent more procedures than average. On the other hand, patient-centered incentive pay appears to have a negative impact on the number of procedures performed, reducing the number of procedures by a statistically significant sixteen percent over the mean.

6.2 Discrete Factor Approximation Models

While the results discussed above do not control for possible selection bias from unobservable characteristics, Table 3 presents results from the discrete factor approximation models with ρ_2 and ρ_3 both fixed at one.[18] These results provide evidence of unobserved heterogeneity bias in the OLS model for time spent with patients but generally confirm the results from the two Tobit models for treatment intensity.

Column (1) presents the results from the discrete factor approximation model for the average time spent with patients. This model confirms the result from the OLS model that productivity incentive pay is associated with a reduction in the average number of minutes a physician spends with each patient, although with a slightly larger impact of approximately two minutes. This finding suggests that productivity incentive pay may be successful in inducing physicians to work harder to see more patients, increasing their overall level of effort. Indeed, if a physician reduces his time spent with each patient from the average of 21 minutes and 47 seconds to 19 minutes and 43 seconds, in an eight hour workday he will see two additional patients. If, however, time spent with the patient is a proxy for quality, productivity incentive pay that increases overall physician effort by decreasing per-patient effort comes at the expense of the quality of medical care received by the patient.

One the other hand, this model does not confirm the effects of patient-centered and practice profiling incentive pay on the average time spent with patients found in the OLS model. Controlling for unobserved heterogeneity, I find that patient-centered

[18] I discuss results from a discrete factor approximation model without this restriction later in this paper.

incentive pay is associated with physicians spending, on average, just over a minutes less with each patient, although this coefficient estimate is imprecisely measured. Further, the results suggest that practice profiling incentive pay has a small, negative, and insignificant impact on the time spent with each patient.

The estimate of the factor loading term in the main equation of this model, ρ_0, is moderately sized at 1.19 although not statistically significant. This may suggest the presence of unobserved heterogeneity leading to bias in the OLS results. Comparing the OLS results with those from the discrete factor approximation model, it appears that the unobserved heterogeneity leads to positive selection bias in the coefficient estimates for patient-centered and practice profiling incentive pay, with less evidence of bias in the coefficient estimate for productivity incentive pay.

Column (3) of Table 3 presents the results from a discrete factor approximation model for the average number of diagnostic and screening services provided. Consistent with those found in the Tobit model without controls for unobserved heterogeneity, the results found here suggest that productivity and patient-centered incentive pay reduce the average number of diagnostic and screening services provided to each patient; however, the size of the effect is much larger although imprecisely estimated.[19] These results also suggest that practice profiling incentive pay increases the average number of diagnostic and screening services provided, again consistent with the results from the Tobit model; however, the magnitude of the coefficient estimate is smaller here and imprecisely

[19] The standard errors are also larger in this model, causing the coefficient estimates to remain statistically insignificant.

estimated. Finally, looking to the estimate of the factor loading term in the main equation, the estimate of the correlation is small and not significantly different from zero, suggesting that while there may be some unobserved heterogeneity in this model, its overall effect should be small.

Finally, the results in column (4) show the effect of incentive pay on the average number of procedures performed on each patient. These results suggest that productivity incentive pay has essentially no impact on the average number of procedures performed. On the other hand, the results indicate that patient-centered incentive pay has a small negative effect on the average number of procedures performed while practice profiling incentive pay has a small positive effect. These effects are smaller in magnitude than, though consistent with, those found in the Tobit model and less precisely estimated. Similar to the model for diagnostic and screening services, I find little evidence of unobserved heterogeneity entering into this model. The estimate of the factor loading term in the main equation is very close to zero and statistically insignificant.

As columns (3) and (4) of Table 3 show, the discrete factor approximation models for treatment intensity confirm the results of the OLS models. Patient-centered incentive pay appears to reduce the number of procedures performed while practice profiling incentive pay seems to increase both the average number of diagnostic and screening services provided and procedures performed. Although the coefficients are less precisely estimated in these models, I find little evidence of unobserved heterogeneity in these two models. Thus one may interpret the OLS estimates as providing evidence of the effect of incentive pay on treatment intensity.

These models also provide estimates of the effects of the explanatory variables on per-patient physician effort. Looking first at the controls for patient characteristics, the results indicate that as the average age of a physician's patients increases, the physician will tend to spend more time with each patient but provide fewer diagnostic and screening services. Physicians with more male patients may spend less time with each patient and provide lower levels of treatment intensity; physicians with more black patients or more Hispanic patients provide more diagnostic and screening services; and physicians with more non-white and non-black patients spend more time with each patient. As the percent of patients expected to pay with public insurance increases, physicians will tend to spend fewer minutes with each patient.

The results also provide information as to how the health of a patient impacts the physician's effort. As the percentage of patients seeing the physician for an acute or chronic problem increases, the physician will, on average, tend decrease treatment intensity; however, the percentage of patients with chronic problems is positively associated with the amount of time spent with each patient. As the percentage of patients seeing the physician for a pre- or post-surgery visit increases, the physician will spend less time with each patient, will perform fewer diagnostic or screening tests, and will order more procedures.

As the proportion of a physician's patients who have been seen before in the medical practice increases, the physician will spend, on average, less time with each patient and will provide a lower level of treatment intensity. When a greater percentage

of patients are seen by their primary care doctors, the physicians tend to spend more time with each patient and provide more diagnostic and screening services.

Observable physician characteristics also impact physician effort. Physicians who are owners in the practice tend to spend about two fewer minutes with the average patient than physicians who are not owners. Physicians who see patients on evenings or weekends tend to order fewer diagnostic and screening services. Surgical and medical specialists tend to provide fewer diagnostic and screening services; however, they do tend to perform more procedures, on average, than primary care physicians. Medical specialists tend to spend more time with their patients.

Finally, I find some evidence that the characteristics of the medical group and the geographic region impact the physician's effort. When a physician's medical practice performs its own lab testing, the physician will spend less time with the average patient but will provide a higher level of treatment intensity. When the medical group uses electronic medical records, physicians tend to spend more time with each patient. When the medical group is located in an MSA, the physician will tend to provide more diagnostic and screening services. Further, physicians who practice in the western United States tend to spend more time with each patient and perform more procedures than those who practice in southern states.

6.3 Analysis by Insurance Type

As mentioned previously, if physicians face different incentive pay for different groups of patients, my assumption of a uniform response to incentive pay across all patients seen by a physician may not be appropriate. This may be most relevant for

patients with private versus public insurance plans. Private insurance plans often have very different reimbursement schedules than Medicare and Medicaid; thus a medical group's revenue function could be decomposed into three independent parts: revenue from private insurance patients, revenue from Medicare patients, and revenue from Medicaid patients. With three independent streams of revenue, a medical group may find it useful to offer physicians incentive pay that varies across private insurance, Medicare, and Medicaid patients. To allow for this possibility, I aggregate the patient-level data up to the physician level in subsamples based on whether the patient is expected to pay primarily with private insurance, Medicare, or Medicaid and repeat the analyses from Table 3 for each subsample. Aggregating the patient-level data by insurance type results in a subsample of 1,756 physicians who saw any private insurance patients during the reporting period, a subsample of 1,408 physicians who saw any Medicare patients during the reporting period, and a subsample of 1,037 physicians who saw any Medicaid patients during the reporting period.[20] Summary statistics for these subsamples are presented in Appendix Table A1.

The results in Table 4 show the impact of incentive pay on physician effort by insurance type. Table 4 presents only the key results for brevity; however, the full set of estimates from the models may be found in Appendix Tables A2, A3, and A4. Columns (1) through (3) of Table 4 show the impact of incentive pay on the average time spent with patients for each of the insurance subsamples. The results indicate that all three types of incentive pay have a stronger, negative impact on the number of minutes

[20] These subsamples are not exclusive.

physicians spend with Medicaid patients than on the time spent with private insurance and Medicare patients. While productivity incentive pay reduces the average amount of time spent with each private insurance and Medicare patient by 2.1 and 2.4 minutes respectively, it reduces the amount of time spent with each Medicaid patient by 4.2 minutes. Patient-centered incentive pay reduces the average amount of time physicians spent with Medicaid patients by over six minutes but has a smaller and statistically insignificant impact on the time spent with private insurance and Medicare patients. Practice profiling incentive pay also has a larger negative impact on the average time spent with Medicaid patients than on the time spent with private insurance or Medicare patients; however, these coefficient estimates are not statistically significant. The estimate of ρ_0 is close to one and statistically insignificant in the models for private insurance and Medicare patients; however, in the model for Medicaid patients, ρ_0 is very large at 5.8 and highly significant, indicating that there is a larger amount of unobserved heterogeneity in this model than in the models for private insurance and Medicare patients.

Columns (4) through (6) of Table 4 show the impact of incentive pay on the average number of diagnostic and screening services provided to each patient. These results suggest that productivity incentive pay and patient-centered incentive pay have approximately the same impact on the average number of diagnostic and screening services provided, regardless of insurance type. Practice profiling incentive pay appears to have a larger, positive impact on the average number of diagnostic and screening services provided to Medicaid patients than to other patients, although this result is not

statistically significant. The estimates of ρ_0 are small and statistically insignificant in these models, suggesting that there is little unobserved heterogeneity in these models.

Columns (7) through (9) of Table 4 show the effect of incentive pay on the average number of procedures performed on each patient by insurance type. The results show that while productivity incentive pay has virtually no impact on the average number of procedures performed on private insurance and Medicare patients, it may slightly increase the number of procedures performed on Medicaid patients, although the estimates are not statistically significant. Patient-centered incentive pay has little if any impact on the average number of procedures performed on private insurance and Medicaid patients, but appears to reduce the number of procedures performed on each Medicare patient by a marginally significant 0.07, a reduction of 37 percent over the mean. Finally, while practice profiling incentive pay does not appear to impact the average number of procedures performed on Medicare patients, it does seem to increase the average number of procedures performed on other patients. For each private insurance patient, practice profiling incentive pay may increase the number of procedures by a marginally significant 0.03, an increase of 18 percent over the mean; and for each Medicaid patient, practice profiling incentive pay may increase the number of procedures by an imprecisely estimated 0.05, an increase of 37 percent over the mean. The estimates of ρ_0 in these models are small and statistically insignificant, indicating little evidence of unobserved heterogeneity.

6.4 Analysis by Health Status

One possible explanation for the results found above is that physicians respond to incentive pay by reducing the amount of time they spend with relatively healthy patients without compromising the care of less healthy patients. If this were the case, the reduction in time spent with patients resulting from incentive pay may not translate into a reduction in the quality of medical care received. In this section, I explore this possibility by aggregating the data into subsamples based on health status. The relatively healthy subsample consists of patients with no chronic conditions, and there are 1,718 physicians who saw patients with no chronic conditions during the reporting period. The less healthy subsample consists of patients with one or more chronic conditions; there are 1,847 physicians who saw patients with one or more chronic conditions in the reporting period. The key results for these subsamples are presented in Table 5; the full results may be found in Appendix Table A6.

As Table 5 shows, incentive pay is associated with a larger negative impact on the time spent with relatively healthy patients than on the time spent with less healthy patients. Productivity incentive pay appears to reduce the average amount of time spent with each patient for both subsamples, although, the magnitude of the result is almost a minute and a half larger for patients without chronic conditions. Patient-centered incentive pay appears to reduce the average time spent with relatively healthy patients by two minutes and fifteen seconds, while the result for less healthy patients is smaller and statistically insignificant. Practice profiling incentive pay also seems to have a larger negative impact on the time spent with patients without chronic conditions, although the

result is not statistically significant for either subsample. Overall, these results suggest that while productivity incentive pay reduces the average amount of time physicians spend with both groups of patients, the effect of patient-centered incentive pay may be limited to relatively healthy patients who may be less likely to see a reduction in quality of medical care from the reduction in time.

6.5 Sensitivity Analysis

As discussed above, identification of the full set of parameters in an unrestricted discrete factor approximation model with ρ_2 and ρ_3 freed up for estimation is theoretically possibly but may not be practically feasible in this application. Indeed, I have found that the model is not fully identified when the dependent variable is censored. Thus, until now, I have discussed results of discrete factor approximation models with ρ_2 and ρ_3 fixed at one. In this section, I will relax that restriction and present results from a discrete factor approximation model for the average time spent with each patient, using the full sample of patient-level data aggregated to the physician level. These results are presented in Table A5 of the appendix.

The results here are very similar to those presented in Table 3. I find that productivity incentive pay is associated with a reduction in the amount of time spent with each patient by one minute and 44 seconds, on average. Patient-centered incentive pay may reduce the average amount of time spent with each patient by about two and a half minutes, although this coefficient estimate is imprecisely measured. Practice profiling incentive pay has a small, negative, and statistically insignificant impact on the amount of

time spent with each patient. The estimate of ρ_0 in this model is large at 2.3 and marginally significant, indicating that unobserved heterogeneity is present in this model.

7. Conclusion

In order to understand the full impact of physicians' incentive pay, including its implications for the cost and quality of medical care, one must know if and how such incentive pay broadly alters physician effort. This paper provides clear evidence to that effect, showing that not only do physicians respond to incentive pay by changing broad measures of per-patient effort, they do so differently for different types of incentive pay and for patients with different types of insurance plans.

I find strong evidence that productivity incentive pay may cause physicians to reduce the average number of minutes spent with each patient. This effect is approximately twice as large for Medicaid patients than for private insurance or Medicare patients and is over sixty percent larger for relatively healthy patients than for those with chronic conditions. Thus productivity incentive pay appears to be a mitigating impact on healthcare costs, reducing the costs associated with physician time; if, however, there are positive benefits associated with a physician spending more time with each patient, productivity incentive pay will result in a decrease in the quality of healthcare.

I also find evidence that patient-centered incentive pay may reduce the amount of time spent with relatively healthy patients, the amount of time spent with Medicaid patients, and the number of procedures performed, especially for Medicare patients. As with productivity incentive pay, these results suggest that patient-centered incentive pay

will help to reduce the costs of medical care, potentially at the expense of reducing the quality of medical care received, particularly for patients with public insurance.

Finally, I find that practice profiling incentive pay may lead to increases in both types of treatment intensity. The effect of practice profiling incentive pay on the average number of diagnostic and screening services provided is strongest for Medicaid patients while the effect on the average number of procedures is strongest for private insurance and Medicaid patients. While productivity and patient-centered incentive pay will help to reduce the costs of medical care, the increase in treatment intensity associated with practice profiling incentive pay will have the opposite effect. Further, if this increase in treatment intensity is not purely an induced demand effect, it will lead to an increase in the quality of medical care as well.

It is important to note the importance of controlling for unobserved heterogeneity in analyses such as this one. Comparing the results from Melichar (2009) and Balkrishnan et al. (2002), one may conclude that unobserved heterogeneity will not bias the estimated impact of financial incentives on physician effort, particularly when effort is measured as the amount of time spent with patients. On the contrary, I find that a failure to account for such unobserved heterogeneity will bias the estimated effect of patient-centered and practice profiling incentives on the average time spent with patients.

Table 1. Average Physician Effort by Incentive Status

	Incentives								
	Productivity			Patient-Centered			Practice Profiling		
	No	Yes	Difference	No	Yes	Difference	No	Yes	Difference
Time spent with patient (Minutes)	22.3	21.1	-1.2*	21.6	22.4	0.8	21.7	22.5	0.7
Number of diagnostic and screening services	3.57	3.59	0.02	3.50	3.83	0.33***	3.54	3.99	0.45**
Number of procedures	0.18	0.17	-0.01	0.18	0.14	-0.04***	0.17	0.18	0.01

Note: Data are from the 2006-2008 NAMCS. There are 1930 observations in the sample. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 2. Regression Results without Controls for Unobserved Heterogeneity: Full Sample

	Dependent Variable		
	Time Spent with Patient (Minutes)	Number of Diagnostic or Screening Services	Number of Procedures
	OLS	Tobit	Tobit
	(1)	(2)	(3)
Productivity incentive pay	-1.732**	-0.041	0.008
	(0.711)	(0.085)	(0.023)
		-0.040	*0.004*
Patient-centered incentive pay	1.056	-0.017	-0.053*
	(0.920)	(0.110)	(0.030)
		-0.017	*-0.028*
Practice profiling incentive pay	1.103	0.269*	0.073*
	(1.249)	(0.150)	(0.040)
		0.259	*0.039*
Average patient age	0.027	0.032***	-0.00033
	(0.021)	(0.003)	(0.001)
		0.030	*-0.0002*
Fraction of male patients	0.012	-0.233	-0.083*
	(1.504)	(0.181)	(0.050)
		-0.224	*-0.044*
Fraction of patients with imputed sex	14.331*	-0.299	0.667***
	(7.495)	(0.907)	(0.233)
		-0.288	*0.355*
Fraction of black patients	0.547	1.258***	-0.044
	(1.764)	(0.212)	(0.060)
		1.211	*-0.023*
Fraction of non-white and non-black patients	6.565***	0.563**	-0.115
	(2.328)	(0.280)	(0.078)
		0.542	*-0.061*
Fraction of patients with imputed race	0.937	-0.165	-0.055
	(1.374)	(0.165)	(0.045)
		-0.159	*-0.029*
Fraction of hispanic patients	-2.168	0.549***	-0.037
	(1.668)	(0.200)	(0.056)
		0.528	*-0.020*
Fraction of patients with imputed ethnicity	-0.944	-0.292*	0.065
	(1.318)	(0.159)	(0.043)
		-0.281	*0.035*

Fraction of patients expected to pay primarily with public insurance	-3.454***	-0.153	0.014
	(1.317)	(0.158)	(0.044)
		-0.148	*0.008*
Fraction of patients expected to pay primarily with means other than private or public insurance	-3.700*	0.04	0.097
	(2.151)	(0.259)	(0.071)
		0.038	*0.052*
Fraction of patients with primary reason for visit being an acute problem	-1.777	-0.669***	0.303***
	(1.867)	(0.224)	(0.062)
		-0.644	*0.161*
Fraction of patients with primary reason for visit being a chronic problem	3.444**	-0.894***	-0.006
	(1.705)	(0.205)	(0.058)
		-0.861	*-0.003*
Fraction of patients with primary reason for visit being a pre-/post-surgery visit	1.003	-0.425	0.477***
	(2.714)	(0.326)	(0.087)
		-0.409	*0.254*
Fraction of patients seen before in the medical practice	-12.566***	-0.489**	-0.06
	(1.925)	(0.232)	(0.064)
		-0.470	*-0.032*
Fraction of patients with imputed values for having been seen before in the practice	-0.827	-0.293	-0.391**
	(5.284)	(0.634)	(0.199)
		-0.281	*-0.208*
Fraction of patients for whom the physician is the primary care doctor	2.737**	0.154	-0.006
	(1.272)	(0.153)	(0.042)
		0.149	*-0.003*
Physician is an MD	0.842	-0.209	-0.082**
	(1.286)	(0.154)	(0.041)
		-0.201	*-0.043*
Physician is an owner in the practice	-2.082***	0.065	0.031
	(0.734)	(0.088)	(0.024)
		0.062	*0.017*
Physician sees patients on evenings or weekends	0.752	-0.145	0.009
	(0.775)	(0.093)	(0.025)
		-0.140	*0.005*
Physician is accepting new patients	-0.294	0.224	-0.065
	(1.655)	(0.199)	(0.054)
		0.215	*-0.034*
Physician specialty is surgical care	0.790	-2.236***	0.238***
	(1.505)	(0.181)	(0.049)
		-2.152	*0.127*

Physician specialty is medical care	5.141***	-1.014***	0.102**
	(1.354)	(0.163)	(0.045)
		-0.976	*0.054*
Medical practice performs its own lab testing	-0.799	0.626***	0.052**
	(0.709)	(0.085)	(0.023)
		0.602	*0.028*
Medical practice has electronic medical records	-0.186	0.117	0.001
	(0.646)	(0.078)	(0.021)
		0.113	*0.001*
The medical practice is a private practice	-0.812	0.188*	0.024
	(0.918)	(0.110)	(0.030)
		0.181	*0.013*
Medical practice is located in an MSA	-1.127	0.375***	-0.038
	(1.058)	(0.127)	(0.034)
		0.361	*-0.020*
Medical practice is located in the northeast	1.120	0.08	-0.014
	(0.919)	(0.110)	(0.030)
		0.077	*-0.007*
Medical practice is located in the midwest	-0.519	0.011	0.025
	(0.840)	(0.101)	(0.027)
		0.010	*0.013*
Medical practice is located in the west	1.660*	-0.088	0.092***
	(0.946)	(0.114)	(0.031)
		-0.085	*0.049*

Note: Data are from the 2006-2008 NAMCS. There are 1,930 observations. Standard errors are in parentheses. Marginal effects on the censored means are in italics * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 3. Restricted Discrete Factor Approximation Model Results: Full Sample

	Dependent Variable		
	Time Spent with Patient (Minutes)	Number of Diagnostic or Screening Services	Number of Procedures
	(1)	(2)	(3)
Productivity incentive pay	-2.045 **	-0.126	0.0003
	(0.958)	(0.156)	(0.039)
		-0.116	*0.0001*
Patient-centered incentive pay	-1.156	-0.152	-0.062
	(1.379)	(0.238)	(0.056)
		-0.139	*-0.018*
Practice profiling incentive pay	-0.019	0.144	0.057
	(1.098)	(0.181)	(0.048)
		0.133	*0.016*
Average patient age	0.023 *	0.030 ***	-0.0004
	(0.014)	(0.003)	(0.001)
		0.027	*-0.0001*
Fraction of male patients	-1.125	-0.286 *	-0.079 *
	(0.828)	(0.155)	(0.044)
		-0.263	*-0.023*
Fraction of patients with imputed sex	0.166	-0.091	0.536 **
	(5.242)	(1.031)	(0.237)
		-0.084	*0.153*
Fraction of black patients	-0.161	1.095 ***	-0.036
	(1.085)	(0.201)	(0.059)
		1.007	*-0.010*
Fraction of non-white and non-black patients	4.038 ***	0.515	-0.093
	(1.332)	(0.353)	(0.086)
		0.474	*-0.027*
Fraction of patients with imputed race	0.241	-0.159	-0.055
	(0.916)	(0.161)	(0.045)
		-0.147	*-0.016*
Fraction of hispanic patients	-1.882	0.399 **	-0.036
	(1.282)	(0.199)	(0.055)
		0.367	*-0.010*
Fraction of patients with imputed ethnicity	-1.896 **	-0.242	0.066
	(0.858)	(0.158)	(0.043)
		-0.223	*0.019*
Fraction of patients expected to pay primarily	-1.632 **	-0.139	-0.002

with public insurance	(0.805)	(0.155)	(0.041)
		-0.128	*-0.001*
Fraction of patients expected to pay primarily with means other than private or public insurance	-1.152	0.019	0.080
	(1.501)	(0.248)	(0.071)
		0.017	*0.023*
Fraction of patients with primary reason for visit being an acute problem	-1.617	-0.668 ***	0.286 ***
	(1.080)	(0.218)	(0.061)
		-0.615	*0.082*
Fraction of patients with primary reason for visit being a chronic problem	2.229 **	-0.918 ***	0.002
	(1.051)	(0.192)	(0.057)
		-0.845	*0.001*
Fraction of patients with primary reason for visit being a pre-/post-surgery visit	-1.931	-0.461	0.399 ***
	(1.514)	(0.300)	(0.073)
		-0.424	*0.114*
Fraction of patients seen before in the medical practice	-9.642 ***	-0.461 **	-0.051
	(1.081)	(0.204)	(0.048)
		-0.424	*-0.015*
Fraction of patients with imputed values for having been seen before in the practice	2.032	-0.184	-0.318
	(3.741)	(0.626)	(0.264)
		-0.169	*-0.091*
Fraction of patients for whom the physician is the primary care doctor	1.008	0.133	-0.007
	(0.762)	(0.147)	(0.040)
		0.123	*-0.002*
Physician is an MD	0.248	-0.224	-0.080 *
	(0.988)	(0.188)	(0.046)
		-0.206	*-0.023*
Physician is an owner in the practice	-2.189 ***	0.036	0.028
	(0.508)	(0.087)	(0.024)
		0.033	*0.008*
Physician sees patients on evenings or weekends	0.339	-0.172 *	0.010
	(0.546)	(0.099)	(0.025)
		-0.158	*0.003*
Physician is accepting new patients	-1.037	0.220	-0.042
	(1.199)	(0.202)	(0.046)
		0.202	*-0.012*
Physician specialty is surgical care	-0.051	-2.227 ***	0.236 ***
	(0.941)	(0.160)	(0.044)
		-2.048	*0.067*
Physician specialty is medical care	4.332 ***	-1.071 ***	0.087 **
	(0.850)	(0.144)	(0.040)

		-0.985	*0.025*
Medical practice performs its own lab testing	-0.804	0.647 ***	0.047 **
	(0.505)	(0.082)	(0.023)
		0.595	*0.014*
Medical practice has electronic medical records	0.867 *	0.114	0.001
	(0.455)	(0.077)	(0.020)
		0.105	*0.000*
The medical practice is a private practice	0.045	0.152	0.014
	(0.593)	(0.108)	(0.029)
		0.140	*0.004*
Medical practice is located in an MSA	-1.086	0.387 ***	-0.028
	(0.698)	(0.128)	(0.032)
		0.356	*-0.008*
Medical practice is located in the northeast	0.928	-0.012	-0.008
	(0.639)	(0.109)	(0.028)
		-0.011	*-0.002*
Medical practice is located in the midwest	-0.516	0.004	0.031
	(0.581)	(0.098)	(0.026)
		0.004	*0.009*
Medical practice is located in the west	1.493 **	-0.079	0.080 ***
	(0.645)	(0.108)	(0.028)
		-0.073	*0.023*
ρ_0	1.188	0.102	0.015
	(0.961)	(0.163)	(0.038)
Point 1	-1.245	-1.244	-1.242
Point 2	1.092 ***	1.102 ***	1.101 ***
	(0.086)	-0.087	-0.087
Weight	0.569 ***	0.564 ***	0.567 ***
	(0.061)	(0.061)	(0.061)
Log Likelihood	-6211.43	-7090.60	-6171.98

Note: Data are from the 2006-2008 NAMCS. There are 1,930 observations. Standard errors are in parentheses. Marginal effects on the censored means are in italics * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 4. Restricted Discrete Factor Approximation Model Results by Type of Insurance

	Time Spent with Patient (Minutes)			Diagnostic and Screening Services			Procedures		
	Private Insurance	Medicare	Medicaid	Private Insurance	Medicare	Medicaid	Private Insurance	Medicare	Medicaid
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Productivity incentive pay	-2.071 **	-2.420 **	-4.218 ***	-0.179	-0.153	-0.148	-0.015	-0.013	0.175
	(0.982)	(1.044)	(0.935)	(0.164)	(0.192)	(0.243)	(0.055)	(0.084)	(0.127)
				-0.164	*-0.137*	*-0.135*	*-0.004*	*-0.003*	*0.040*
Patient-centered incentive pay	-0.539	-0.698	-6.101 ***	-0.137	-0.161	-0.214	-0.044	-0.293 *	0.001
	(1.398)	(1.638)	(1.023)	(0.245)	(0.342)	(0.337)	(0.081)	(0.150)	(0.204)
				-0.125	*-0.144*	*-0.194*	*-0.013*	*-0.072*	*0.000*
Practice profiling incentive pay	-0.234	-0.571	-1.719	-0.018	-0.046	0.208	0.113 *	0.019	0.236
	(1.167)	(1.386)	(1.097)	(0.206)	(0.263)	(0.269)	(0.066)	(0.113)	(0.187)
				-0.017	*-0.041*	*0.189*	*0.033*	*0.005*	*0.054*
ρ_0	0.990	1.149	5.795 ***	0.172	0.150	0.098	0.017	0.124	-0.181
	(0.966)	(1.103)	(0.653)	(0.164)	(0.223)	(0.234)	(0.054)	(0.092)	(0.137)
Log Likelihood	-5641.64	-4518.29	-3343.97	-6460.93	-5198.86	-3825.88	-5257.96	-3876.61	-2596.82
Observations	1756	1408	1037	1756	1408	1037	1756	1408	1037

Note: Data are from the 2006-2008 NAMCS. Standard errors are in parentheses. Marginal effects on the censored means are in italics. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 5. Restricted Discrete Factor Approximation Model Results by Health Status

	Time Spent with Patient	
	No Chronic Conditions	One or More Chronic Conditions
	(1)	(2)
Productivity incentive pay	-3.559 ***	-2.147 **
	(0.890)	(0.993)
Patient-centered incentive pay	-2.249 **	-1.248
	(1.011)	(1.312)
Practice profiling incentive pay	-1.317	-0.416
	(1.022)	(1.169)
Log Likelihood	-5477.27	-5910.12
Observations	1718	1847

Note: Data are from the 2006-2008 NAMCS. Standard errors are in parentheses. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

CHAPTER III

EXITING TANF IN SOUTH CAROLINA AFTER THE DEFICIT REDUCTION ACT

1. Introduction

The Deficit Reduction Act of 2005 (DRA) made a number of changes to the Temporary Assistance for Needy Families (TANF) program with the goal of increasing the economic self-sufficiency of welfare recipients. As part of these changes, the DRA standardized across states the work and work readiness activities that satisfy the TANF work requirement. Being that states could previously define their own work activities, this policy change had a unique impact on the TANF program in each state. For South Carolina in particular, it meant narrowing the set of activities that may be used to satisfy the TANF work requirement.

This essay uses administrative records from South Carolina to examine the impact of the DRA's change in allowable work activities on the likelihood of welfare recipients to exit the state's TANF program. To accomplish this task, I employ descriptive and multivariate event-history techniques with a difference-in-difference estimation framework to identify the impact of the DRA on the overall likelihood of exit from the TANF program in South Carolina, on the likelihood of exit through three different paths, and on the duration of benefit receipt. I find that in narrowing the list of allowable work and work readiness activities, the DRA had a very different impact on black recipients in South Carolina than on non-blacks, leading overall to longer durations of TANF

participation for blacks but shorter durations for non-blacks. The DRA also had a differential impact on exits due to employment for black and non-black recipients.

Over the past two decades, welfare reform has sought to strengthen work incentives, and a breadth of literature has looked at the effects of these incentives on welfare recipients (see, e.g., Blank (2002), Moffitt (2003), and Grogger and Karoly (2005) for reviews). This project adds to the literature in three ways. First, it is one of very few studies on the impact of welfare reform since 2000, let alone on the DRA. As Ribar and Wolff (2013) show, the TANF caseload declined by almost 50 percent between 2000 and 2008, and the economic circumstances of those who leave TANF worsened. The impact of welfare reform must be considered in light of these changes. By focusing this study around the DRA legislation of 2005, this project accomplishes that task.

The second way in which this analysis contributes to the literature is by improving upon the methodology of prior studies. While many studies have examined the impact of work requirements on welfare dependency, the success of such research has been hampered by the fact that a multitude of reforms were implemented all at around the same time and many with opposing effects. In this study, I make use of South Carolina's administrative dataset to isolate the impact of one particular reform: the DRA's standardization of work and work readiness activities. In addition to being able to isolate this single reform, I am also able to identify treatment and control groups of welfare recipients, allowing estimation of the impact of the reform under a difference-in-difference framework.

Finally, little research has been done on the impact of work requirements on the paths through which recipients exit the TANF program. The third contribution of this paper is a competing risks analysis of the impact of the DRA reform on the paths to exit from South Carolina's TANF program. In particular, I examine three exit paths: through employment, through increases in other income, and through administrative reasons.

2. TANF and the DRA

The welfare reauthorization component of the DRA was enacted largely in response to policy makers' dissatisfaction with two components of the TANF program's work requirement: the work participation standard and the activities that were considered "work." First, the TANF program's work participation standard states that fifty percent of all recipient households must participate in work activities, and should a state fail to meet its work participation standard, the state risks losing its federal TANF funding. In an effort to encourage states to move recipients off of welfare and into work, each state is given a credit towards the work participation standard for declines in its welfare caseload. This way, those who leave TANF are counted as workers for the purpose of meeting the work participation requirement. During the first seven years of the TANF program, the caseload was declining at such a rate that the work participation standard was not binding for many states after the credits for caseload reduction were factored in. Policymakers became dissatisfied with the ineffectuality of the work participation standard. Second, a 2005 GAO report identified a leniency and lack of consistency across states in the types of activities that could count toward meeting the federal work requirement. For example, some states allowed activities such as bed rest, exercise, and helping a friend with errands

to count as work regardless of whether or not such activities contribute to self-sufficiency (GAO, 2005). Policymakers wished to have more control over these activities in the hopes of gearing them towards activities that are more successful at promoting self-sufficiency.

The Deficit Reduction Act of 2005 attempted to remedy these issues. When the DRA came into effect on October 1, 2006, it strengthened the TANF work rules through three main channels: (1) by defining the activities that satisfy the work requirement; (2) by adjusting the work participation standard based on declines in a state's caseload from its 2005 level rather than from its 1995 level; and (3) by requiring that families in separate State programs meet the work requirements (Federal Register, 2006). This project examines the impact of the first change, the definition of work activities, in the context of South Carolina's TANF program. Because South Carolina had previously set its own definition of work activities, the definition set by the DRA effectively changed the set of activities that satisfy the work requirement for welfare recipients in South Carolina.

2.1. The Family Independence Program

The TANF program in South Carolina is known as the Family Independence (FI) program (SC DSS, 2009). The primary goal of the FI program is to assist families in becoming self-sufficient. The program consists of three main elements geared towards achieving this goal: (1) work requirements; (2) FI cash benefits; and (3) support services.

The first element, the FI work requirement, is designed to promote self-sufficiency through full-time employment. All FI recipients who are work-eligible,

meaning they are deemed able to work by the federal government, must agree to a Family Plan and an Employability Plan detailing the actions the recipient will take to become gainfully employed. If a recipient fails to cooperate in developing a Family Plan, she risks losing her TANF benefits in South Carolina. The ultimate goal of the work requirement is for the recipient to find a job within 24 months that offers a wage rate above the minimum wage, health benefits, potential for advancement, and convenience of location.

While all work-eligible recipients must enter into a Family Plan and an Employability Plan, some of these recipients are exempt from mandatory participation in the work requirement. The majority of women who are exempt are single parents with a child under one year of age. A recipient may also be exempt if she is providing care for a disabled child who attends school, if she is a victim of domestic violence, or if she lacks childcare or transportation.

The second element of the FI program is the cash benefit. This is the money paid to recipients in the FI program. Support services, the third element of the FI program, are non-monetary forms of assistance, such as childcare and transportation, which are provided to remove some barriers to employment and self-sufficiency.

2.2 The DRA in South Carolina

Because each state had implemented its own TANF program uniquely tailored to fit its needs, the DRA was likely to impact each state differently. For TANF recipients in South Carolina, the impact was felt almost exclusively in the change in allowable work activities (those activities that would satisfy the work requirement). Prior to the DRA,

South Carolina's TANF case managers were given the authority to assign recipients to the work activities that they felt were most appropriate for the recipient. After the DRA, much of that authority was removed. Case managers could now only assign activities that the federal government had defined as work or work readiness and were required to provide documentation verifying that the requirement had been met. These new rules caused case managers to change how they assigned work activities. For example, South Carolina could no longer count time spent preparing for and traveling to a job interview as a work-related activity, and the time spent in the job interview needed to be verified. The difficulty of verification combined with the small amount of countable time led case managers to decrease the frequency with which they assigned recipients to job search, instead assigning recipients to other activities that may not lead as quickly to gainful employment.

While the DRA's definition of allowable work activities caused South Carolina to make changes to its TANF program, adjusting the work participation standard did not. Between fiscal years 2005 and 2006, South Carolina's TANF caseload declined by approximately two percent, reducing South Carolina's work participation standard from 50 percent to 48 percent. South Carolina received an additional credit for spending in excess of the required amount on the TANF program in fiscal year 2006, further reducing the work participation standard to 29 percent (US DHHS, 2012). This standard requires that 29 percent of all families in South Carolina's TANF program participate in work activities in order for South Carolina to receive federal TANF funding. Yet from 2002 through 2009, the percentage of TANF recipients engaged in work activities was near or

above 50 percent, indicating South Carolina should have had little difficulty meeting its work participation standard and thus should have not been constrained by the second change made by the DRA (U.S. DHHS, 2003, 2004, 2006, 2007a, 2007b, 2009a, 2011d, 2011e).[1] Further, South Carolina was able to bypass the third change made by the DRA by paying benefits to families in separate state programs with non-TANF related funds. Thus, the DRA had its primary impact on the TANF program in South Carolina through the change in the allowable work activities.

3. Welfare Reform Literature

An abundance of literature has attempted to measure the impacts of welfare reforms, including those specifically related to work requirements (see Blank (2002), Moffitt (2003), and Grogger and Karoly (2005) for reviews). A drawback of many of these studies, however, is the difficulty involved in trying to isolate the impact of any single reform. When states enacted reforms, they generally changed multiple aspects of their welfare programs simultaneously, making identification of the effect of any single reform difficult. Further, these studies tend to be sensitive to the time period under analysis, resulting in often conflicting and counterintuitive findings (Blank, 2002; Moffitt, 1999).

There have been some studies that have successfully examined the impact of work requirements on welfare use and the employment of welfare recipients. Moffitt (1996) finds that work requirements increased the rate of exit from the welfare program. Blank

[1] The work participation standard for fiscal years 2008 and 2009 are not yet available; however, the TANF caseload in South Carolina was declining over this period, suggesting that the work participation standard was at its peak in 2007 (U.S. DHHS, 2009b, 2010c).

(2002) finds, in her review of the literature, that mandatory work programs significantly increased employment and decreased welfare use. Further, Grogger and Karoly (2005) find that work-related programs increased in the employment rate of welfare recipients by 5.6 percentage points. On the other hand, Fang and Keane (2004) highlight the difference between leaving welfare and working. They find that work requirements accounted for 57 percent of the decrease in welfare participation from 1993 through 2002 but only 17 percent of the increase in work participation. These results suggest that work requirements were more successful at getting women to exit welfare but perhaps less successful at inducing work. Indeed, close to one quarter of welfare leavers during this time period were not employed (Fang and Keane, 2004).

Given the finding by Frogner, Moffitt, and Ribar (2009) that "work pays", one might think that if work requirements can increase employment, they will lead to increased self-sufficiency of welfare recipients. Unfortunately, this has not always been found to be the case. As Blank (2002) points out in her review, the increases in income resulting from work requirements tend to be completely offset by the loss in welfare benefits. Further, in a review of a number of early studies examining the impact of work requirements on the earned income of welfare recipients, Moffitt (1992) finds that the increases in earnings were not large enough to greatly impact self-sufficiency.

When it comes to increasing earnings, however, studies have found that not all work requirements are created equal. Programs that focused on job search or labor market attachment saw increases in earnings in the short-run while those that focused on human capital development took longer to experience those gains (Blank, 2002). In the

long-run, on the other hand, intensive training programs appear to have more persistent impacts on earnings than short-term job search and readiness programs (Dyke et al., 2006). Bloom and Michalopoulos (2001) examine the results from twenty programs that randomly assigned welfare recipients into job search programs, education and training programs, or both. They find that both types of programs led to an increase in the earnings of program recipients; however, the programs that had the most success were those that combined job search with education and training. Ultimately, Bloom and Michalopoulos conclude that programs tailored to the individual needs of the recipient will be more successful at increasing earnings, and thus promoting self-sufficiency, than a "one-size-fits-all" approach. This conclusion suggests that, in standardizing the work requirement and removing the authority of the case managers in each state to assign work requirements, the DRA may have actually impeded self-sufficiency rather than promoted it.

While the research discussed above has shown that work requirements led to decreased welfare participation, there is more uncertainty regarding how those decreases took place. Are recipients exiting due to increases in earned income, or are they being sanctioned off the welfare rolls due to noncompliance? Fang and Keane (2004) do show that there is a discrepancy between welfare exits and work participation associated with work requirements, but they do not examine the paths through which those exits take place. Some studies have examined the impact of education on the paths to exit from welfare. Harris (1993) finds that more educated welfare recipients were more likely to find a job that paid enough so that they could immediately exit welfare, while recipients

with prior work experience were more likely to work while receiving welfare benefits and eventually become self-sufficient. Similarly, Blank and Ruggles (1996) find that higher levels of education increased the likelihood of exiting welfare due to income and other reasons not associated with family composition.

In contrast to the bevy of literature examining the impact of welfare reforms in the twentieth century, a limited number of studies have considered the impact of the DRA on welfare recipients. A GAO (2010) report does find that the proportion of welfare recipients engaged in work activities decreased slightly nationwide following implementation of the DRA. The report also finds that many states, including South Carolina, chose to fund certain low-income families with separate state funds, removing these families completely from the TANF program and excluding them from the calculation of the work percentage. Shifting recipients to programs that are funded with separate states dollars gives the appearance that the number of families receiving assistance has decreased (Pavetti et al., 2009).

This study adds to the literature by identifying the impact of a single reform, namely the DRA's definition of allowable of work activities, on the likelihood of exit from the TANF program in South Carolina and on the paths to exit. As such, I am able to isolate the effect of this single reform rather than being faced with the task of attempting to estimate the impact of multiple, simultaneous, and possibly opposing, reforms. This study also contributes to the literature by examining the effect of work requirements on the paths to exit from the TANF program. Further, I use a difference-in-difference

estimator to capture the impact of the DRA's definition of work activities on the exit behavior of TANF recipients in South Carolina.

4. Data

The data I use in this analysis are a full extract of administrative records from the South Carolina FI program. The period of analysis begins in January 2002, when the federal PRWORA work requirements came into full effect, and ends in September 2009, just prior to the changes implemented by the American Recovery and Reinvestment Act of 2009. The administrative records contain information on the household characteristics, demographics, and benefit receipt of each FI program participant. I transform the data so that they describe spells of cash benefit receipt, where each spell represents a string of consecutive months of benefit receipt, with one observation per household per month of positive cash benefit receipt. The spell begins when the household enters the FI program and receives benefits, and it ends when the household exits the FI program and is thus no longer receiving cash benefits.

Given the history and demography of South Carolina, it seems plausible that blacks may have been differentially impacted by the DRA reform. The demographic characteristics available in the administrative data allow me to identify the race of the each FI program participant, and I use this information to disaggregate the data into two subsamples. The subsample of black households consists of those with a head of household who is black. The subsample of non-black households similarly consists of those households whose head is not black.

4.1 Treatment and Control Groups

The difference-in-difference analysis requires identification of both a treatment and a control group. I identify non-exempt work-eligible households as my treatment group, as these are the households that were impacted by the DRA's definition of work activities in South Carolina. Non-exempt work-eligible households are those with a head of household who is an FI participant who is deemed work-eligible and not exempted from South Carolina's work requirement. My control group consists of exempt work-eligible households whose youngest child is at least three months old. This group provides a natural control group for the analysis because the heads of these households have been deemed able to work just like those in the treatment group; yet, as I described in Section 2.1, South Carolina has exempted them from the work requirement, making them not subject to the DRA's definition of work activities.

4.2 Sample Selection

In order to focus on only those cases that may be impacted by the DRA, I restrict the dataset to cases with single heads of households who are work-eligible FI program participants throughout the entire spell and who are not minor parents. This eliminates cases where only the children in the household receive benefits, where there are two parents in the household receiving benefits, and where the head of household is disabled.

In order to limit the data inconsistencies and coding errors in the dataset, I smooth the data by removing one-month spells of participation and non-participation. Finally, I omit left-censored spells from the dataset as I do not observe the true durations of benefit receipt for these households. The final dataset consists of 38,330 spells of FI benefit

receipt. For each spell of benefit receipt, the administrative data provide information on demographic and household characteristics such as the number of people in the household and the age, race, sex, and education level of each person in the household. The administrative data also provide the reason for exit for completed spells. For each month of a spell, I merge in county-level unemployment data from the Bureau of Labor Statistics. There are 21,622 households receiving FI benefits in the sample, and approximately seventeen percent of households experience multiple spells during the study period.

For households that exit the FI program, the reason for exit in the administrative data provides me with information on the path to exit for my analysis. First, I identify households that exit due to an increase in earned income as having employment exits. Second, I identify households that exit due to an increase in unearned or total income as having other income exits. Third, I identify households that exit for all other reasons as having administrative exits.[2] Twenty-six percent of spells end due to employment, five percent end due to increases in other income, and 64 percent end for administrative reasons. The remainder of the spells are right-censored.

4.3 Sample Statistics

As the first column of Table 1 shows, the median spell length for the full sample is approximately five months. Almost all of the households have female heads. Sixty-four percent of heads of households are high school graduates and 75 percent are black.

[2] Other reasons for exit include such things as reaching the time limit, being sanctioned for failing to comply with program rules, and voluntarily withdrawing from the program.

At the start of the spell, the average age of the head of household is almost 26 years, the average age of the oldest child in the household is 4.6 years, and the average age of the youngest child in the household is almost three years.

Columns (2) through (5) of Table 6 provide summary statistics for non-exempt and exempt work-eligible households before and after implementation of the DRA. Eighty-five percent of the monthly observations in the sample belong to the group of non-exempt work-eligible households. As the table shows, all three types of exits were slightly less common for both groups after the DRA was implemented. This is likely due to the fact that spells that start in the latter period are more likely to be right-censored. Interestingly, though, the median spell length for both groups fell following implementation of the DRA. Both groups also had a higher incidence of high school graduates in the post-DRA period. Heads of non-exempt households tended to be older after the DRA was implemented, while the children of both groups tended to be older in the post-DRA period. There was also a slightly lower incidence of female headship in both groups following implementation of the DRA, although the overwhelming majority of household heads were still female.

5. Methods

This analysis employs a difference-in-difference approach to estimate the impact of the DRA's definition of work activities on the exit behavior of TANF recipients in South Carolina. The difference-in-difference estimator compares the changes in exit behavior after implementation of the DRA of TANF recipients who are and are not

subject to the DRA's regulations. I perform all analyses in this essay for the full sample and for the two subsamples disaggregated by race.

The difference-in-difference estimator relies heavily on the appropriateness of the control group, as it attributes any unobserved differences between the treatment and control groups in the change in the likelihood of exit after implementation of the DRA to the Deficit Reduction Act itself. Exempt work-eligible households should be very similar to the nonexempt work-eligible households in terms of their likelihood of exit from the TANF program and their labor market prospects except for the fact that the exempt work-eligible households are not subject to the work requirement and thus not affected by the DRA's definition of work activities. The majority of households in the control group are exempted because they consist of a single parent with a child under the age of one. If household heads with very young children are less likely to work than those with older children, this will compromise the appropriateness of the control group; however, as Klerman and Leibowitz (1994) show, most women who are working when their child is one year old had returned to work well before that, within three months of childbirth. This finding suggests that the one-year post-childbirth mark is not a critical node in women's work behavior, and women with a child between three months and one year of age should comprise an adequate control group for women with a child over one year of age. Further, FI recipients qualify for subsidized childcare, helping to remove this barrier to work for women with young children. As I will show, while exempt work-eligible households are slower to the leave the TANF program in South Carolina than non-

exempt work-eligible households, their patterns of exit are similar prior to implementation of the DRA.

5.1 Descriptive Approach

In analyzing the impact of the DRA's definition of work activities on the exit behavior of TANF recipients in South Carolina, I take both a descriptive and multivariate approach. For my descriptive analysis, I estimate Kaplan-Meier hazard rates for both non-exempt and exempt work-eligible households before and after the DRA was implemented. The hazard rate is the probability that a household exits the TANF program in a particular month given that the household is still receiving benefits in that month. Using these Kaplan-Meier hazard estimates, I construct difference-in-difference estimates of the hazard rate. I calculate the difference-in-difference estimate as the difference between two differences. First, I calculate the non-exempt difference as the difference between the hazard rate after and before DRA implementation for non-exempt work-eligible households. I calculate the exempt difference in the same way for exempt work-eligible households. Then, the difference-in-difference estimate is the difference between the non-exempt difference and the exempt difference. A positive difference-in-difference estimate suggests that the DRA's definition of work activities increased the likelihood of exiting the TANF program in South Carolina, while a negative difference-in-difference estimate suggests that the reform reduced the likelihood of exit.

5.2 Multivariate Approach

For my multivariate analyses, I employ discrete-time logistic hazard and competing risks models under a difference-in-difference framework to estimate the effect

of the DRA's definition of work activities on the likelihood of exit and on the path to exit from South Carolina's TANF program. I construct the DRA difference-in-difference policy variables by interacting a dummy variable for non-exempt work-eligible households and a dummy variable for the period in which the DRA is in effect in my data, from October 2006 through September 2009, with the baseline hazard. These interactions result in a set of duration-specific controls for non-exempt work-eligible households from October 2006 through September 2009, allowing me to identify the effect of the DRA at various durations on the group of welfare recipients affected by the legislation, as compared to those who were not targeted by the reform. Thirty-seven percent of observations in the sample belong to non-exempt work-eligible households in the post-DRA period.

The discrete-time logistic hazard model estimates the effect of the DRA on the overall likelihood of exit from the TANF program. I model the hazard rate, $h(t)$, as

$$h(t) = \frac{\exp(\beta_0 T_t + \beta_1 (NonExempt_t \cdot T_t) + \beta_2 (DRA_t \cdot T_t) + \beta_3 ((NonExempt_t \cdot DRA_t) \cdot T_t) + \delta X_t + \lambda \Psi_t + \alpha)}{1 + \exp(\beta_0 T_t + \beta_1 (NonExempt_t \cdot T_t) + \beta_2 (DRA_t \cdot T_t) + \beta_3 ((NonExempt_t \cdot DRA_t) \cdot T_t) + \delta X_t + \lambda \Psi_t + \alpha)} \quad (6)$$

(Allison, 1984). $NonExempt_t$ is a dummy variable for households with non-exempt work-eligible heads, and DRA_t is a dummy variable for the period in which the DRA is in effect. T_t is a vector of controls for the baseline hazard and consists of a linear spline for months one through four, a linear spline for months five through twelve, a linear spline for months thirteen and up, and dummy variables at months twelve and twenty-four to capture key dates in South Carolina's TANF program. The key coefficients of interest in

this model are found in the vector β_3. These coefficients capture the effect of the DRA's definition of work activities on non-exempt work-eligible households' likelihood of exit relative to that of exempt work-eligible households at each duration. X_t is a vector of all other observed exogenous variables at time t such as household size, household demographic characteristics, and county-level unemployment, and Ψ_t is a linear monthly time trend to control for unobserved changes over time that may impact the likelihood of exit.

The presence of unobserved heterogeneity may lead to spurious negative duration dependence, which would bias my estimator if not controlled for. To account for this, I estimate each model with a random effect, α, to control for unobserved heterogeneity.

The competing risks model examines the effect of the DRA on the three paths to exit: employment, other income, and program administration. Given these $J = 3$ paths to exit from the TANF program, I estimate, using a competing risks model, the impact of the DRA on each path-specific hazard rate (Allison, 1984). I model the hazard rate, $h_j(t)$, for each path j as

$$h_j(t) = \frac{\exp(\beta_{0j}T_t + \beta_{1j}(NonExempt_t \cdot T_t) + \beta_{2j}(DRA_t \cdot T_t) + \beta_{3j}((NonExempt_t \cdot DRA_t) \cdot T_t) + \delta_j X_t + \lambda_j \Psi_t + \alpha)}{1 + \exp(\beta_{0j}T_t + \beta_{1j}(NonExempt_t \cdot T_t) + \beta_{2j}(DRA_t \cdot T_t) + \beta_{3j}((NonExempt_t \cdot DRA_t) \cdot T_t) + \delta_j X_t + \lambda_j \Psi_t + \alpha)} \quad (7)$$

where $j = (1, 2, 3)$ represents the three paths to exit from South Carolina's TANF program and all other parameters and variables are as in the discrete-time logistic hazard model described above.

A positive estimate of any parameter in β_{3j} indicates that the DRA increased the likelihood that a household will exit the TANF program through path j at that duration, while a negative estimate of a parameter in β_{3j} indicates that the DRA reduced the likelihood of exit through path j at that duration. The baseline hazard pattern is also as described in the discrete-time logistic hazard model except that the equations for employment exits and other income exits do not include the dummy variables for month twenty-four due to a lack of exits through these paths in the twenty-fourth month.

I estimate the logistic hazard and competing risks models using the full sample of administrative records as well as with a subsample made up of black households and a subsample made up of non-black households. I separate the sample by race to account for the possibility that blacks may be differentially impacted by the DRA reform. This is of particular concern given the history and demography of South Carolina.

6. Results

6.1 Descriptive Results

Figure 1 presents Kaplan-Meier hazard estimates for the full sample of non-exempt and exempt work-eligible households before and after the DRA came into effect. The conditional likelihood of both groups to exit the TANF program in South Carolina was generally larger after the DRA was implemented. The likelihood of exit grew through the first 4 months. After the fourth month, it leveled off until the twenty-fourth month, reflecting the state's twenty-four month time limit. The spikes in the hazard estimates at 24 months represent a large likelihood of exit; however, very few households actually face this probability. As Figure 2 shows, the majority of spells are short, ending

within the first five months for non-exempt work-eligible households and within the first seven months for exempt work-eligible households. In fact, only ten percent of all spells survive past twelve months, and only 3.5 percent survive past eighteen months.

Figure 1 also provides some insight into the appropriateness of using exempt work-eligible households as a control group for non-exempt work-eligible households in this analysis. Panels A and C of Figure 1 show the Kaplan-Meier hazard estimates for the conditional likelihood of exit of both non-exempt and exempt work-eligible households prior to implementation of the DRA. Through the ninth month of TANF benefit receipt, although the hazard rate is lower for exempt work-eligible households than for non-exempt households, the trend in the hazard rate is very similar for both groups: it increases through the first four months, decreases in months five and six, has a small spike in month seven before further decreasing in months eight and nine. After nine months, the hazard rate of non-exempt work-eligible households is rather steady, with spikes in months twelve and twenty-four. The hazard rate of exempt work-eligible households is more volatile due to the small number of exempt households with long spells; however, the hazard rate is fairly similar to that of non-exempt households at longer spell lengths. The similarity of the trends in the hazard rate for non-exempt and exempt work-eligible households suggests that households with exempt work-eligible heads may provide a good comparison group for households with non-exempt work-eligible heads; however, the difference in the levels of the hazard rates across the two groups, to the extent that it cannot be accounted for with observable characteristics,

indicates that there may be some limitation to the strength of exempt work-eligible households as a control group in this analysis.

Figure 3 presents the Kaplan-Meier hazard estimates for the black subsample. The hazard rate for both non-exempt and exempt black households was larger after the DRA was implemented than before, particularly during the first four months of a spell. This suggests that both groups of black recipients were more likely to exit the TANF program in South Carolina early on the in the spell after the DRA was implemented.

Figure 4 presents the Kaplan-Meier hazard estimates for the non-black subsample. For non-exempt non-black households, the hazard rate was larger after implementation of the DRA, particularly during the first eleven months of a spell. Exempt non-black households did not experience the same increase in their hazard rate following the DRA's implementation. For exempt non-black households, the conditional likelihood of exit appears somewhat similar, if not smaller, after the DRA was implemented.

Table 7 displays the Kaplan-Meier hazard rates presented in Figure 1 along with the calculation of the difference-in-difference estimate. At most durations, the hazard rate experienced a small increase after implementation of the DRA for both non-exempt and exempt households. This indicates that the likelihood of exiting the TANF program in South Carolina was generally larger after the DRA was implemented than before for both groups. The difference-in-difference estimates are generally small and not significantly different from zero, suggesting that the impact of the DRA's change in allowable work activities was limited; however, exempt households often saw a larger increase in the likelihood of exit than did non-exempt households, leading to a negative

difference-in-difference estimate at many durations. In particular, at durations of two months and sixteen months, the difference-in-difference estimate is negative and significantly different from zero, indicating that at these durations, the DRA's definition of work activities may have decreased the likelihood of exit from the TANF program in South Carolina for non-exempt work-eligible households. At durations of 21 and 24 months, the difference-in-difference estimate is larger in magnitude, positive, and significantly different from zero, suggesting that the reform may have increased the likelihood of exit at these durations.

Table 8 presents the difference-in-difference estimates calculated from the Kaplan-Meier hazards by exit path. For employment exits, the DRA's definition of allowable work activities appears to have had a negative impact on the likelihood of exit early on in spells and a positive impact at durations of 21 months. The reform also appears to have had a negative impact on the likelihood of exit due to increases in other income at durations of four, fifteen, sixteen, and seventeen months. On the other hand, while the DRA appears to have reduced the likelihood of exit for administrative reasons at durations of sixteen months, at other durations, it appears to have had a positive effect. In particular, the difference-in-difference estimates are positive and significantly different from zero at durations of four, nine, and 21 months, suggesting that the DRA's definition of work activities may have increased the likelihood of administrative exit at these durations.

Table 9 presents the difference-in-difference estimates calculated from the Kaplan-Meier hazards for the black subsample. These results suggest that the DRA's

definition of work activities reduced the likelihood of exit early on in spells. For all exits regardless of exit path and for employment exits, the difference-in-difference estimate is negative and marginally significant at durations of two months. For other income exits, the difference-in-difference estimate is negative and statistically significant at durations of four months; and for administrative exits, the difference-in-difference estimate is negative and marginally significant at durations of seven months. On the other hand, the reform appears to have increased the likelihood of employment exits for black households at durations of seven months.

Table 10 presents the difference-in-difference estimates calculated from the Kaplan-Meier hazards for the non-black subsample. For all exits and administrative exits, the DRA's change in allowable work activities appears to have had a strong, positive impact on the likelihood of exit in the fourth month. The reform also appears to have had a negative, marginally significant impact on employment exits at durations of two months and a strong positive impact on other income exits at longer durations.

These descriptive results provide mixed evidence of the effect of the DRA's change in allowable work activities on the likelihood of exit from the TANF program in South Carolina. For the full sample, the reform generally appears to have reduced the likelihood of exit for any reason early in spells and increased it at longer durations. This finding holds for exits due to employment. The likelihood of other income exits appears to have been more negatively impacted by implementation of the DRA, while it appears that the reform had more of a positive impact on administrative exits. When I divide the sample by race, I find that the negative impact of the DRA's definition of work activities

holds for blacks but not for non-blacks. In fact, I find strong positive impacts of the reform for the non-black subsample. Of course, these results do not control for observable factors; however, they do suggest that the DRA's definition of work activities failed to have a strong, uniform impact on the hazard rate and that it may have differentially impacted blacks and non-blacks.

6.2 Multivariate Results

The first column of Table 11 presents results from a discrete-time logistic hazard model estimated on the full sample controlling for unobserved heterogeneity with a random effect. The coefficients on the interactions between non-exempt work eligible households in the DRA period and the duration dependence variables are small and statistically insignificant. These results suggest that the DRA's change in allowable work activities had little impact on the likelihood of recipients to exit the TANF program in South Carolina. However, dividing the sample by race, I find that the reform did differentially impact blacks and non-blacks. The results for blacks, which are presented in column 2, suggest that in the first four months of a spell, the DRA's definition of work activities reduced the likelihood of exit from South Carolina's TANF program. On the other hand, non-blacks appear to have benefited from an increased likelihood of exiting the TANF program in the first four months in response to the reform. Given that most spells end within five months, it seems reasonable that the impact of the reform would be felt at these shorter durations.

Household composition appears to be an important determinant of exit behavior, regardless of race. Households that were headed by a woman or by someone who is

black were less likely to leave South Carolina's TANF program, while households that were headed by a high school graduate were more likely to exit the program. As the household head aged, the household became less likely to exit South Carolina's TANF program. Having children of any age reduced the likelihood of exit, but as the oldest child aged, that likelihood increased. The likelihood of exiting the program increased as households accumulated previous months of benefit receipt and decreased as the unemployment rate rose. The estimates also provide information on the baseline duration dependence pattern. In all three models, the likelihood of exit increased with the duration of the spell and spiked at 24 months.

Table 12 presents the coefficient estimates from the competing risks models for the full sample, and Tables 13 and 14 present the results for the black and non-black subsamples, respectively. For the full sample, the results indicate that the DRA's definition of work activities reduced the likelihood of exit due to employment in the first four months of the spell but had little impact on other income and administrative exits. The results from the model estimated on the black subsample also indicate that the reform reduced the likelihood of employment exit in the first four months of a spell but had little effect on exits through the other two paths. In contrast, the model for non-blacks indicates that while the reform did not impact employment exits for this group, it did increase the likelihood of administrative exit in the first four months and exit due to other income at durations longer than 4 months.

Having a head of household who is female or black reduced the hazard for all three types of exit, while a head of household who is a high school graduate had an

increased likelihood of exit through employment or other income but a smaller administrative exit hazard. The hazard rate for all three types of exit decreased with the age of the household head, decreased with the number of children in the household, and increased with the age of the oldest child in the household.

7. Simulations

While the estimates from the logistic hazard and competing risks models provide information as to how the DRA's definition of work activities impacted the hazard rate of TANF recipients, these models do not tell us directly what effect the reform had on spells of TANF receipt in South Carolina. I conduct simulations in order to determine how the reform impacted the length of spells. Using a simulated dataset based on the actual administrative data and the estimates from the logistic hazard and competing risks models, I am able to show that overall and for spells ending in employment, the DRA's definition of work activities lengthened the spells of black recipients in South Carolina while shortening those of non-black recipients. For spells ending due to increases in other income and due to administrative reasons, the reform appears to have increased spell lengths for both groups.

I begin constructing the simulated dataset using the information from the first month of each actual spell that started on or after October 2006 (when the DRA was first implemented), replicating each spell five times. I advance each spell 36 months or up to September 2009, whichever comes first. For each month of the simulated spell, I merge in actual county-level unemployment data from the BLS. I also update the age information for each member of the household every twelve months. Otherwise, each

month of a simulated spell contains the household and demographic information from the first month of an actual spell. I assign unobserved heterogeneity to each spell with a random draw from a normal distribution with mean zero and standard deviation estimated from the corresponding logistic hazard or competing risks model.

I conduct simulations for the full sample, for the subsample made up of black recipients, and for the subsample made up of non-black recipients. For each of the three samples, I conduct a baseline simulation, a simulation of a scenario where the DRA is never implemented (the "No DRA" simulation), and a simulation of a scenario where everyone is subject to the DRA (the "All DRA" simulation). For all three simulations and all three samples, I simulate the likelihood of any exit from the TANF program in South Carolina corresponding to the logistic hazard model, and I simulate the likelihood of exit through employment, other income, and administration corresponding to the competing risks model. In all, I conduct eighteen simulations.

7.1 Simulation of All Exits

For the baseline simulation of any exit, I use the coefficient estimates from the logistic hazard model to calculate the probability of exit for each simulated household in each month and compare it to a random draw from a uniform distribution. If the calculated probability of exit is greater than the random draw, I simulate exit in that month and the spell ends. Otherwise, the simulated spell continues. Figure 5 shows the Kaplan-Meier hazard estimates for the full sample baseline simulation for two groups: non-exempt households and exempt households. As expected, the hazard estimates are very similar to those of the actual data in Panels B and D of Figure 1. Table 10 shows

statistics for spell length. The statistics for the full sample baseline simulation show slightly longer spell durations than the actual data. Median spell length is five months for the baseline simulation. Sixty-five percent of spells in the baseline simulation last at least four months; 39 percent last at least six months; and fourteen percent last at least twelve months.

To show what would have happened to spell length had the DRA never been implemented, I construct a "No DRA" simulation for the full sample where the variable for the DRA is always zero. The resulting spell length statistics are reported in the third row of Table 10. Had the DRA's change in allowable work activities never come into effect, spells would have been slightly shorter. The median spell length in this simulation is four months and 28 days, two days shorter than in the baseline simulation. The fraction of spells lasting at least four, six, and twelve months is also slightly smaller. Sixty-four percent of spells last at least four months, 38 percent last at least six months, and thirteen percent last at least twelve months.

I also conduct a full-sample simulation to examine what spells would have looked like if everyone had been subject to the reform brought on by the DRA. The results from this "All DRA" simulation are presented in the fourth row of Table 15. Had the reform been in effect for all recipients, spells of TANF receipt in South Carolina would not have been much different. The median spell length in this simulation is five months and one day, only one day longer than in the baseline simulation. The fractions of spells lasting at least four, six, and twelve months are nearly identical to those from the baseline simulation.

The remainder of Table 15 repeats the simulations by race. For the black subsample, again I find that the baseline simulation overestimates median spell length as compared to the actual data. In the "No DRA" simulation, I find that median spell length would have been six days shorter and the fractions of spells lasting four, six, and twelve months would have been smaller had the DRA not been implemented for blacks. In the simulation where everyone is subject to the DRA's change in allowable work activities, I find that median spell length is 1.5 days longer than in the baseline simulation, and the fractions of spells lasting six and twelve months are slightly larger.

For the non-black subsample, I continue to find that the baseline simulation overestimates median spell length. In contrast to the simulation results for the black sample, I find that median spell length for non-blacks would have been longer had the DRA's reform not been implemented. In the "No DRA" simulation, median spell length is four months and 28 days, as opposed to four months and eighteen days in the baseline. The fractions of spells lasting four, six, and twelve months are also larger than in the baseline simulation for non-blacks. Had all non-blacks been subject to the DRA's change in allowable work activities, the results indicate that median spell length would have been slightly shorter than baseline at four months and seventeen days. The fractions of spells lasting four, six, and twelve months would have also been slightly smaller.

7.2 Simulations of Exit Paths

To simulate exits from TANF by exit path, I use the coefficient estimates from the competing risks model to calculate the probability of exit for each exit path and compare those probabilities to a random draw from a uniform distribution. If the probability of

exiting due to employment is greater than the random draw, I simulate exit in that month through employment. If the random draw is greater than the probability of exit due to employment but less than the sum of the probabilities of exit due to employment and other income, I simulate exit in that month through other income. If the random draw is greater than the sum of the previous two probabilities but less than the sum of all three probabilities, I simulate exit in that month through administrative reasons. Otherwise, the simulated spell continues.

Table 16 presents results from the simulations of spells that end due to employment. For the full sample, I find that had the DRA's change in allowable work activities not come into effect, the median spell length of those spells that end in employment would have been approximately two days shorter. However, the "All DRA" scenario does not seem to have greatly impacted median spell length of employment exit spells for the full sample.

For blacks, I find that had the reform not come into effect, the median length of spells that end due to employment would have been shorter and the fractions of spells lasting four, six, and twelve months would have been smaller. The median length of employment exit spells is four months and 29 days in the baseline simulation and four months and 23 days in the "No DRA" simulation. Had the DRA's reform been in effect for all work-eligible households, the median spell length would have been one day longer than in the baseline simulation. For the non-black sample, the median spell length in the "No DRA" scenario of spells that end due to employment is 22 days longer than the baseline, whereas the median spell length in the "All DRA" simulation is three days

shorter than the baseline. These results indicate that the DRA's definition of work activities slowed down exit through employment for blacks yet increased the speed at which non-blacks exit TANF due to employment in South Carolina.

Table 17 presents results from the simulations of other income exits. For the full sample, I find that the DRA's change in allowable work activities led to longer lengths of spells that end due to increases in other income. Had the reform not come into effect, the median spell ending due to other income would have been almost five days longer, and had everyone been subject to the reform, the median spell ending due to other income would have been almost one day longer. For the black subsample, the results from the simulations indicate that the DRA's definition of work activities may have slightly lengthened spells that end due to increases in other income. The results for non-blacks are similar yet more pronounced. For non-blacks, I find that had everyone been subject to the DRA's reform, the median spell ending due to other income would have been seven days longer than baseline.

Table 18 presents results from the simulations of administrative exits for the full sample and by race. For the full sample, I find that the spells ending for administrative reasons would have been slighter shorter had the DRA's change in allowable work activities not come into effect and slightly longer had everyone been subject to the reform. For the black subsample, I find a similar result. The median spell length in the "No DRA" simulation is almost three days shorter than in the baseline, while the median spell length in the "All DRA" simulation is three days longer than baseline. For non-

blacks, I find very little impact of the reform on the median length of spells that end for administrative reasons.

8. Conclusion

The results found here suggest that the DRA's definition of work activities reduced the hazard rate for blacks but increased the hazard for non-blacks in the first four months of a spell. For blacks, this decrease came from a decrease in the likelihood of employment exits. For non-blacks, the increase in the hazard rate in the first four months of a spell came from an increase in the likelihood of administrative exits. Results from simulations suggest that had the DRA's reform not come into effect, the median spell would have been one fifth of a month shorter for blacks and three tenths of a month longer for non-blacks.

If the DRA's definition of work activities is thought to have promoted self-sufficiency among TANF recipients, the results found here for blacks are counterintuitive; however, it is possible that the reform failed to promote self-sufficiency among this group in South Carolina. The DRA imposed a "one-size-fits-all" approach to the activities that satisfy the work requirement, and states were no longer able to tailor their work activities to best fit the needs of their recipient populations. Prior research has shown that this is not the way to promote self-sufficiency (Bloom and Michalopoulos, 2001). In a state such as South Carolina, where the TANF program is less accommodating at the outset, this could have an especially deleterious effect, particularly for the most disadvantaged recipients. Case managers, who assign work activities to TANF recipients in South Carolina and who may be the individuals most familiar with

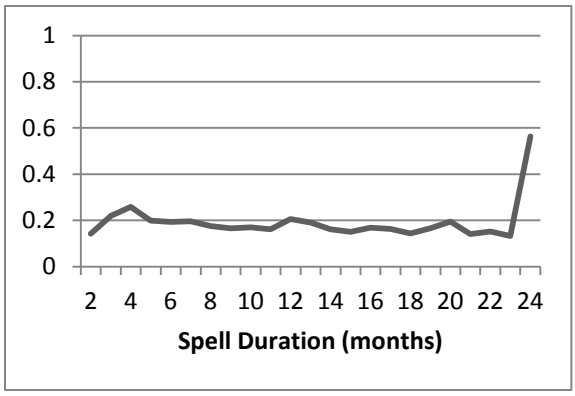
the needs of TANF recipients, have lost authority over what can count as work for each recipient under the DRA and thus may be less successful in helping a TANF recipient achieve self-sufficiency. If black recipients in South Carolina rely more heavily on the assistance of case managers in achieving self-sufficiency than do their non-black counterparts, then it may come as no surprise that the reform enacted by the DRA differentially impacted blacks and non-blacks.

A further consideration has to do with the definitions of the allowable work activities themselves. Some activities, such as job search, were changed in such a way that it made it much harder for a recipient in South Carolina to satisfy the work requirement with that activity. For instance prior to the DRA, a recipient in South Carolina who contacted ten employers in search of a job in a given week was said to have satisfied the thirty-hour work requirement. Once the DRA's change in allowable work activities came into effect, the amount of time allowed for each employer contact was reduced so that a TANF recipient may have to contact thirty or forty employers in order to fulfill the work requirement with job search alone. For TANF recipients living in rural or more disadvantaged areas of South Carolina, there may not be that many job openings for a recipient to apply to; thus case managers were forced to assign recipients to other work activities. If TANF recipients were being redirected into activities that are less successful at promoting self-sufficiency in order for the recipient to satisfy the work requirement, then the results of this study may come as no surprise.

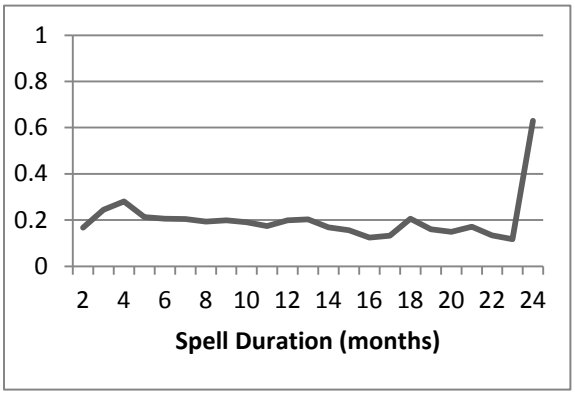
Finally, while these results provide evidence that the DRA's definition of work activities may have failed to increase the self-sufficiency of black TANF recipients in

South Carolina while promoting that of whites, one should be cautious before extrapolating the results to TANF populations in other states. As mentioned previously, each state's TANF program is unique, and South Carolina is no exception. While the DRA differentially impacted black and non-black recipients in South Carolina, it did not necessarily have the same effect in all states across the country. Indeed, it is possible that the DRA had a positive impact on all recipients in some states; more research is needed to determine this. Yet for South Carolina, this study suggests that the DRA's definition of work and work readiness activities may have reduced the self-sufficiency of some TANF recipients rather than increased it.

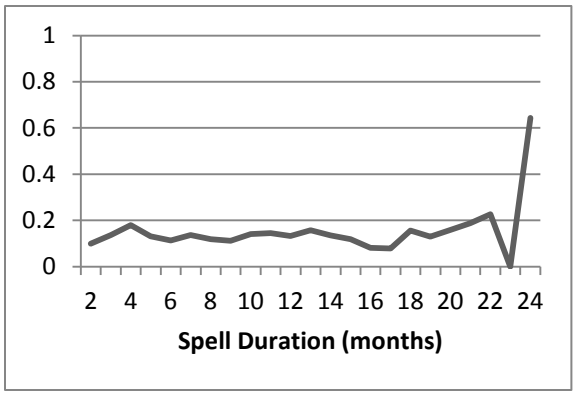
Figure 1. Kaplan-Meier Hazard Estimates from the South Carolina Administrative Data



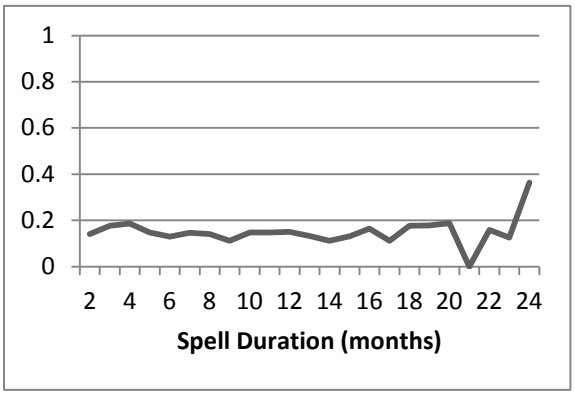
Panel A: Non-exempt work-eligible households prior to the DRA



Panel B: Non-exempt work-eligible households after the DRA



Panel C: Exempt work-eligible households prior to the DRA



Panel D: Exempt work-eligible households after the DRA

Figure 2. Survival Rates Estimated from the South Carolina Administrative Data

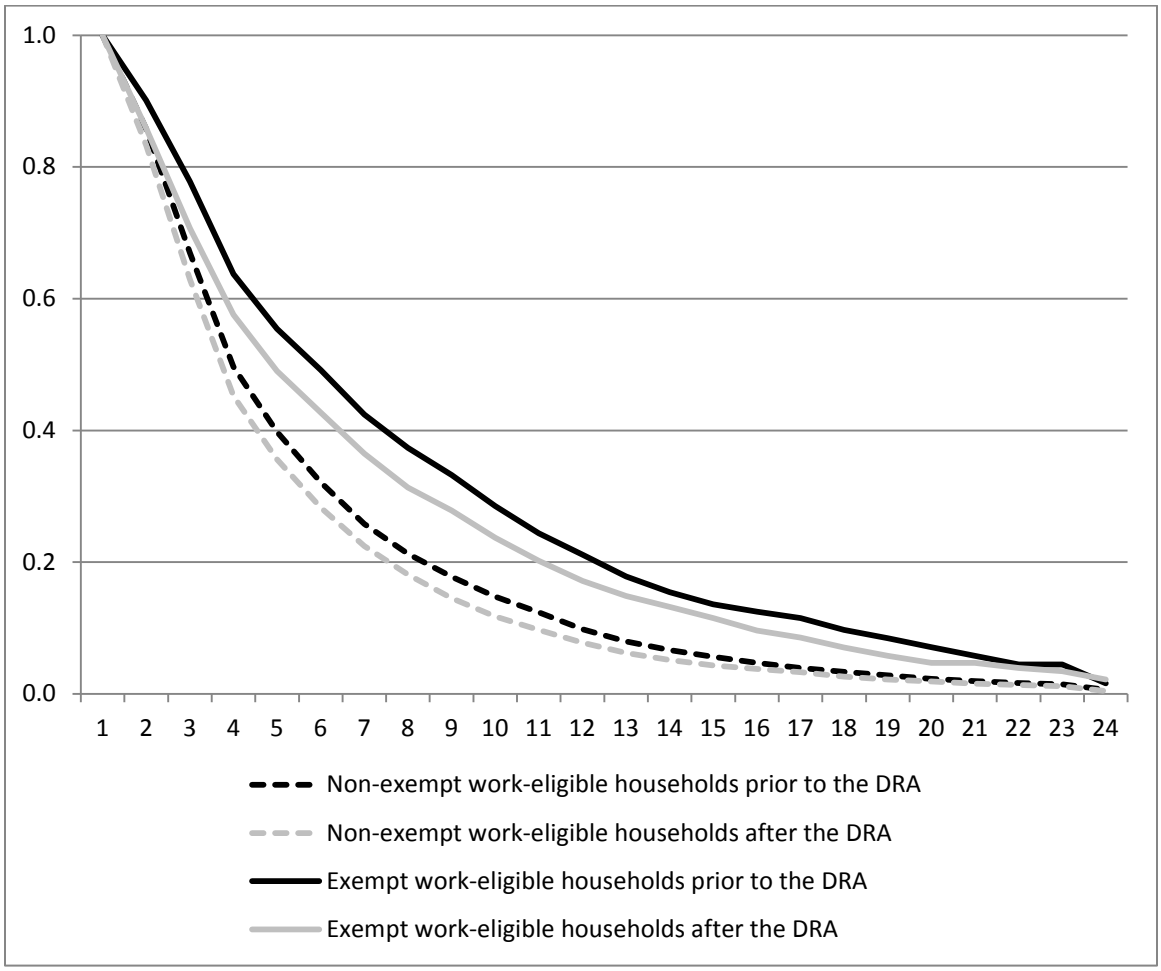
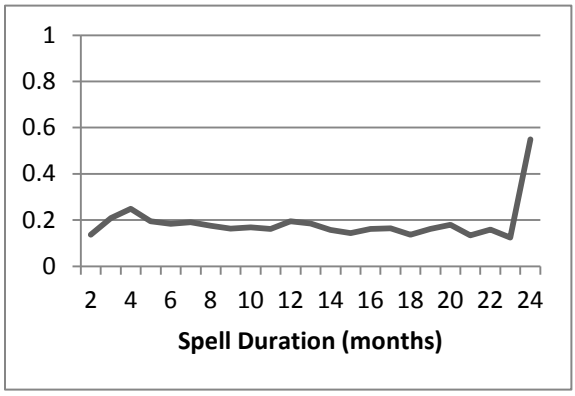
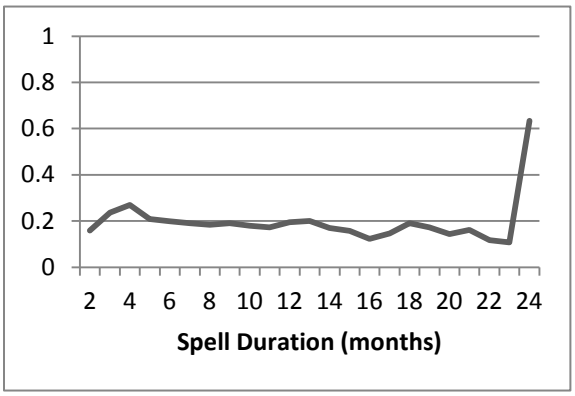


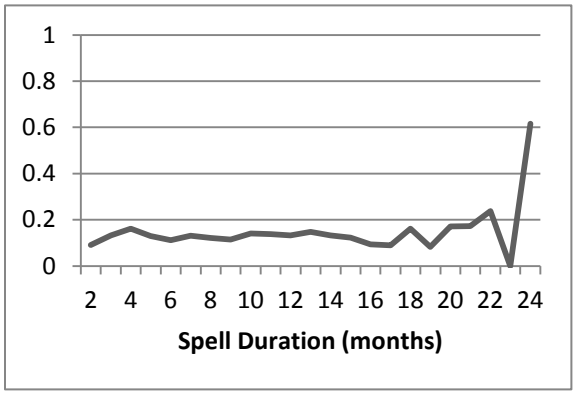
Figure 3. Kaplan-Meier Hazard Estimates for the Black Subsample



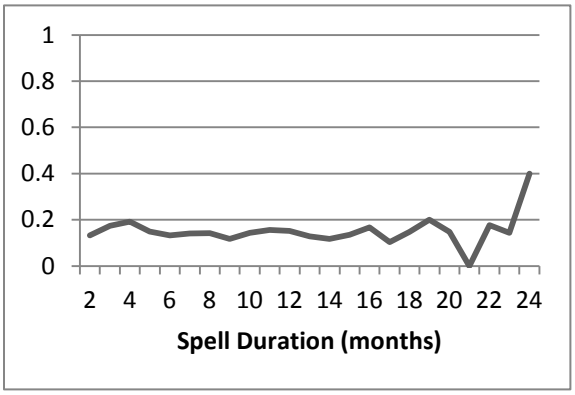
Panel A: Non-exempt work-eligible households prior to the DRA



Panel B: Non-exempt work-eligible households after the DRA

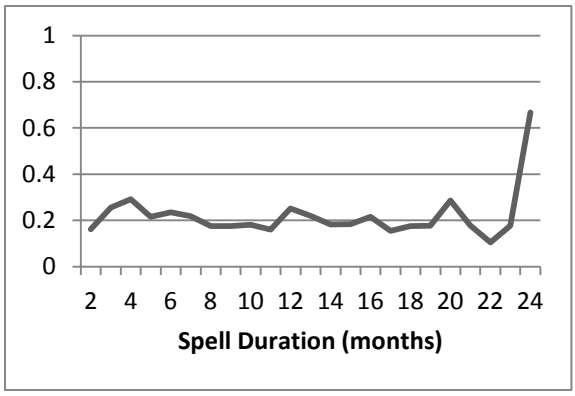


Panel C: Exempt work-eligible households prior to the DRA

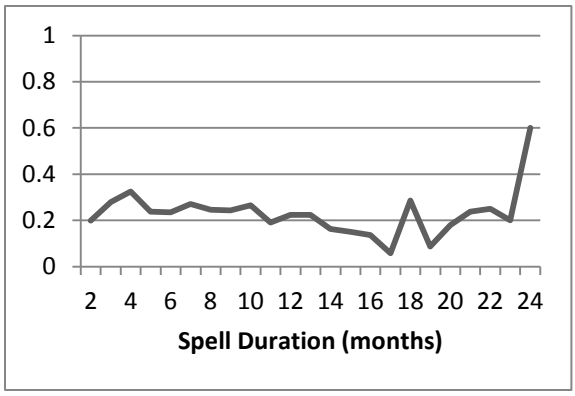


Panel D: Exempt work-eligible households after the DRA

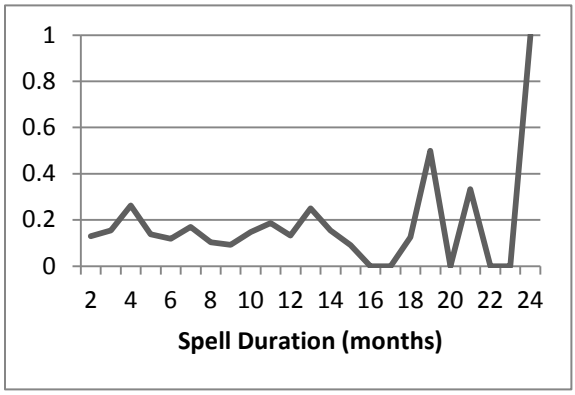
Figure 4. Kaplan-Meier Hazard Estimates for the Non-Black Subsample



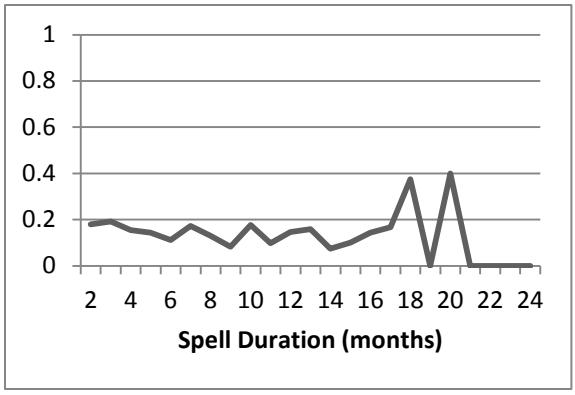
Panel A: Non-exempt work-eligible households prior to the DRA



Panel B: Non-exempt work-eligible households after the DRA

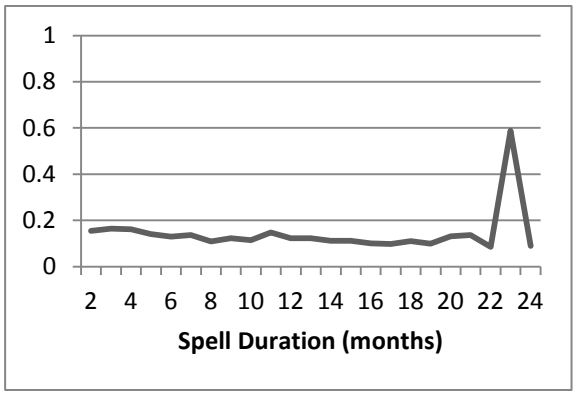


Panel C: Exempt work-eligible households prior to the DRA

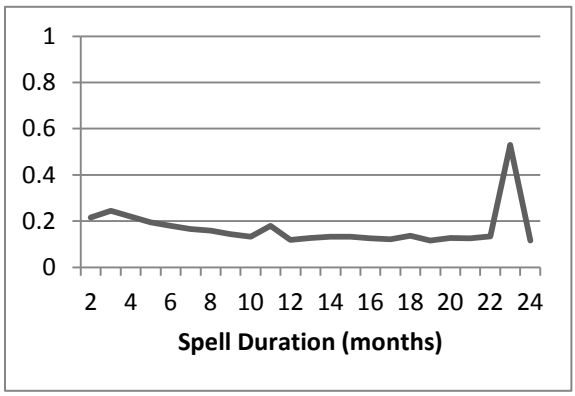


Panel D: Exempt work-eligible households after the DRA

Figure 5. Kaplan-Meier Hazard Estimates from the Baseline Simulation



Panel A: Non-exempt work-eligible households after the DRA



Panel B: Exempt work-eligible households after the DRA

Table 6. Summary Statistics

Characteristics of the Spells	Full Sample	Non-exempt Households Prior to DRA	Non-exempt Households After DRA	Exempt Households Prior to DRA	Exempt Households After DRA
	(1)	(2)	(3)	(4)	(5)
Employment Exit	0.260	0.277	0.227	0.298	0.282
	(0.002)	(0.003)	(0.003)	(0.009)	(0.009)
Other Income Exit	0.054	0.059	0.051	0.05	0.034
	(0.001)	(0.002)	(0.002)	(0.004)	(0.003)
Administrative Exit	0.642	0.664	0.625	0.652	0.574
	(0.002)	(0.003)	(0.004)	(0.009)	(0.009)
Age of Head of Household at Start of Spell	25.8	25.9	26.3	24.4	24.3
	(0.030)	(0.043)	(0.053)	(0.092)	(0.089)
Age of Oldest Child at Start of Spell	4.6	4.4	5.2	3.2	3.7
	(0.016)	(0.020)	(0.030)	(0.053)	(0.063)
Age of Youngest Child at Start of Spell	2.9	3	3.4	1.5	1.6
	(0.014)	(0.017)	(0.025)	(0.041)	(0.043)
Median Spell Length (months)	4.98	4.98	4.74	6.86	5.93
Number of Spells	38,330	18,530	14,333	2,681	2,786
Characteristics of the Households					
Female Head of Household	0.978	0.978	0.971	0.995	0.992
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
High School Graduate	0.635	0.607	0.636	0.693	0.779
	(0.003)	(0.004)	(0.006)	(0.011)	(0.011)
Black	0.746	0.742	0.733	0.789	0.788
	(0.003)	(0.004)	(0.006)	(0.009)	(0.010)
Number of Households	21,622	11,797	6,428	1,874	1,523
Time-varying Characteristics					

Non-Exempt * DRA Time Period	0.367	0	1	0	0
	(0.001)				
DRA Time Period	0.452	0	1	0	1
	(0.001)				
Non-Exempt	0.849	1	1	0	0
	(0.001)				
Age of Head of Household	25.7	25.8	26	24.7	24.5
	(0.013)	(0.019)	(0.023)	(0.042)	(0.035)
Age of Oldest Child	4.6	4.4	5.1	3.6	4.1
	(0.007)	(0.009)	(0.013)	(0.024)	(0.025)
Age of Youngest Child	2.8	3.0	3.2	1.4	1.5
	(0.006)	(0.008)	(0.011)	(0.018)	(0.017)
Number of Children Age 0-2	0.669	0.586	0.635	0.999	1.035
	(0.001)	(0.002)	(0.002)	(0.006)	(0.005)
Number of Children Age 3-5	0.583	0.625	0.559	0.5	0.516
	(0.001)	(0.002)	(0.002)	(0.005)	(0.005)
Number of Children Age 6-11	0.418	0.383	0.501	0.296	0.354
	(0.002)	(0.002)	(0.003)	(0.005)	(0.005)
Number of Children Age 12-14	0.027	0.01	0.056	0.004	0.022
	(<0.000)	(<0.000)	(0.001)	(0.001)	(0.001)
Number of Children Age 15-17	0.004	0.004	0.006	0.001	0.001
	(<0.000)	(<0.000)	(<0.000)	(<0.000)	(<0.000)
Total months of benefit receipt	9.7	8.6	11.2	7.7	10.7
	(0.016)	(0.020)	(0.030)	(0.050)	(0.059)
Unemployment Rate	0.078	0.073	0.086	0.071	0.081
	(<0.000)	(<0.000)	(<0.000)	(<0.000)	(<0.000)
Number of Observations	192,814	92,825	70,842	12,916	16,231

Note: Standard errors are in parentheses.

Table 7. Kaplan-Meier Hazard Rates and Difference-in-Difference Estimates for All Exits

	Non-Exempt Work-Eligible Households		Exempt Work-Eligible Households		Non-Exempt Difference	Exempt Difference	Difference-in-Difference
	Before DRA	After DRA	Before DRA	After DRA			
Duration	(1)	(2)	(3)	(4)	(2) - (1)	(4) - (3)	[(2) - (1)] - [(4) - (3)]
2 months	0.142	0.167	0.099	0.141	0.025	0.042	-0.017 *
3 months	0.22	0.245	0.137	0.177	0.025	0.04	-0.015
4 months	0.258	0.281	0.18	0.186	0.023	0.006	0.017
5 months	0.199	0.213	0.131	0.148	0.014	0.017	-0.003
6 months	0.193	0.206	0.113	0.129	0.013	0.016	-0.003
7 months	0.196	0.205	0.137	0.146	0.009	0.009	0
8 months	0.175	0.193	0.119	0.141	0.018	0.022	-0.004
9 months	0.165	0.199	0.111	0.112	0.034	0.001	0.033
10 months	0.17	0.191	0.141	0.148	0.021	0.007	0.014
11 months	0.162	0.174	0.145	0.148	0.012	0.003	0.009
12 months	0.206	0.199	0.133	0.151	-0.007	0.018	-0.025
13 months	0.191	0.203	0.157	0.132	0.012	-0.025	0.037
14 months	0.162	0.169	0.135	0.111	0.007	-0.024	0.031
15 months	0.151	0.156	0.119	0.131	0.005	0.012	-0.007
16 months	0.169	0.124	0.081	0.164	-0.045	0.083	-0.128 ***
17 months	0.163	0.133	0.078	0.111	-0.03	0.033	-0.063
18 months	0.143	0.206	0.156	0.177	0.063	0.021	0.042
19 months	0.165	0.16	0.13	0.178	-0.005	0.048	-0.053
20 months	0.195	0.149	0.159	0.188	-0.046	0.029	-0.075
21 months	0.14	0.171	0.188	0	0.031	-0.188	0.219 ***
22 months	0.152	0.134	0.227	0.158	-0.018	-0.069	0.051
23 months	0.132	0.117	0	0.125	-0.015	0.125	-0.14

24 months	0.563	0.629	0.643	0.364	0.066	-0.279	0.345 *

Note: There are a total of 38,676 spells of TANF benefit receipt. There are no exits in the first month. * indicates significantly different from zero at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 8. Difference-in-Difference Estimates of Kaplan-Meier Hazard Rates by Exit Path

Duration	Employment Exits	Other Income Exits	Administrative Exits
2 months	-0.017 ***	0.001	-0.001
3 months	-0.021 ***	0.004	0.001
4 months	-0.014	-0.007 **	0.038 ***
5 months	-0.01	0.002	0.003
6 months	-0.007	0.006	-0.006
7 months	0.017	0.005	-0.022
8 months	-0.012	-0.007	0.015
9 months	0.002	-0.002	0.033 *
10 months	-0.018	0.004	0.028
11 months	-0.01	0.004	0.015
12 months	-0.02	-0.002	-0.003
13 months	-0.008	0.004	0.04
14 months	-0.008	-0.005	0.043
15 months	-0.008	-0.011 *	0.012
16 months	-0.022	-0.015 *	-0.091 **
17 months	-0.008	-0.004 *	-0.051
18 months	0.029	0.011	0
19 months	-0.037	-0.002	-0.015
20 months	-0.06	0.016	-0.031
21 months	0.086 *	0.006	0.127 *
22 months	0.02	0.001	0.03
23 months	-0.058	-0.008	-0.075
24 months	0.078	0.071	0.194

Note: There are a total of 38,676 spells of TANF benefit receipt. There are no exits in the first month. * indicates significantly different from zero at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 9. Difference-in-Difference Estimates of Kaplan-Meier Hazard Rates for the Black Subsample

Duration	All Exits	Employment Exits	Other Income Exits	Administrative Exits
2 months	-0.020*	-0.036*	0.015	-0.016
3 months	-0.014	-0.038	0.039	-0.008
4 months	-0.010	-0.037	-0.223**	0.030
5 months	-0.007	-0.023	0.040	0.009
6 months	-0.005	0.001	0.162	0.000
7 months	-0.01	0.106**	0.059	-0.040*
8 months	-0.013	0.010	-0.206	0.004
9 months	0.025	0.079	-0.036	0.046
10 months	0.007	0.002	0.176	0.025
11 months	-0.008	-0.048	0.315	0.011
12 months	-0.019	-0.038	-0.170	0.008
13 months	0.034	-0.034	0.543	0.074
14 months	0.029	-0.094	-0.018	0.090*
15 months	0	-0.077	-0.313	0.058
16 months	-0.113**	-0.201		-0.087
17 months	-0.031	0.053		-0.028
18 months	0.067	0.241		0.035
19 months	-0.107	-0.368		-0.069
20 months	-0.013	-1.015***		0.016
21 months	0.199**			0.193**
22 months	0.02			0.001
23 months	-0.159			-0.106
24 months	0.299			0.202

Note: There are a total of 30,183 spells of TANF benefit receipt among black households. There are no exits in the first month. Estimates are not calculated for employment exits past 20 months and for other income estimates past 15 months due to a lack of exits at these longer durations. * indicates significantly different from zero at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 10. Difference-in-Difference Estimates of Kaplan-Meier Hazard Rates for the Non-Black Subsample

Duration	All Exits	Employment Exits	Other Income Exits	Administrative Exits
2 months	-0.012	-0.072*	0.080	0.006
3 months	-0.014	-0.107	-0.059	0.025
4 months	0.141***	0.069	0.130	0.185***
5 months	0.018	0.034	-0.109	0.028
6 months	0.007	0.022	-0.020	0.023
7 months	0.05	0.008	0.649***	0.036
8 months	0.044	-0.047	-0.048	0.105*
9 months	0.078	0.010	0.185	0.116*
10 months	0.056	-0.042	0.086	0.089
11 months	0.118	0.113	0.127	0.119
12 months	-0.041	-0.328**	0.958***	-0.002
13 months	0.097	-0.222		0.179
14 months	0.061	0.064		0.138
15 months	-0.041	-0.159		0.030
16 months	-0.222**	-0.030		-0.312*
17 months	-0.265**	-0.360*		-0.179
18 months	-0.139	-0.417		0.059
19 months	0.409*	-0.161		0.564**
20 months	-0.507**	-1.214***		-0.295
21 months	0.393			0.492
22 months	0.145			
23 months	0.024			
24 months	0.933***			

Note: There are a total of 8,493 spells of TANF benefit receipt among non-black households. There are no exits in the first month. Estimates are not calculated for employment exits past 20 months, for other income estimates past 12 months, and for administrative exits past 21 months due to a lack of exits at these longer durations. * indicates significantly different from zero at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 11. Discrete-time Logistic Hazard Estimates

	Full Sample	Black Subsample	Non-black Subsample
	(1)	(2)	(3)
Female	-1.367 ***	-1.746 ***	-1.033 ***
	(0.051)	(0.069)	(0.088)
High School Graduate	0.117 ***	0.123 ***	0.057 *
	(0.016)	(0.019)	(0.032)
Black	-0.369 ***		
	(0.020)		
Age of Household Head	-0.012 ***	-0.009 ***	-0.024 ***
	(0.001)	(0.002)	(0.003)
Age of Oldest Child	0.056 ***	0.056 ***	0.063 ***
	(0.006)	(0.007)	(0.013)
Age of Youngest Child	-0.042 ***	-0.050 ***	-0.027 **
	(0.006)	(0.007)	(0.012)
Number of Children Age 0-2	-0.302 ***	-0.335 ***	-0.262 ***
	(0.020)	(0.023)	(0.044)
Number of Children Age 3-5	-0.210 ***	-0.203 ***	-0.283 ***
	(0.016)	(0.018)	(0.037)
Number of Children Age 6-11	-0.132 ***	-0.114 ***	-0.205 ***
	(0.021)	(0.023)	(0.046)
Number of Children Age 12-14	-0.096 *	-0.075	-0.138
	(0.050)	(0.058)	(0.103)
Number of Children Age 15-17	-0.209 *	-0.380 ***	0.091
	(0.109)	(0.130)	(0.225)
Total Months of Benefit Receipt	0.024 ***	0.024 ***	0.018 ***
	(0.001)	(0.001)	(0.003)
Unemployment Rate	-5.151 ***	-5.445 ***	-4.766 ***
	(0.303)	(0.347)	(0.642)
Baseline Duration Dependence			
Spline 1-4 Months	0.134 ***	0.126 ***	0.143 ***
	(0.020)	(0.023)	(0.045)
Spline 5-12 Months	0.015	0.033 *	-0.052
	(0.016)	(0.018)	(0.040)
Spline 13+ Months	0.055 **	0.050 *	0.100
	(0.027)	(0.028)	(0.080)
Dummy 12th Month	0.129	0.094	0.278
	(0.216)	(0.233)	(0.591)
Dummy 24th Month	2.377 ***	2.276 ***	2.031
	(0.592)	(0.603)	(1.543)

Non-exempt work-eligible households interacted with			
Spline 1-4 Months	0.155 ***	0.165 ***	0.119 ***
	(0.012)	(0.014)	(0.025)
Spline 5-12 Months	-0.037 **	-0.046 ***	0.009
	(0.015)	(0.017)	(0.038)
Spline 13+ Months	-0.004	-0.001	-0.038
	(0.027)	(0.029)	(0.079)
Dummy 12th Month	0.249	0.228	0.305
	(0.224)	(0.243)	(0.603)
Dummy 24th Month	-0.638	-0.562	0.034
	(0.609)	(0.622)	(1.507)
Post-DRA period interacted with			
Spline 1-4 Months	0.043 ***	0.066 ***	-0.039
	(0.016)	(0.018)	(0.036)
Spline 5-12 Months	-0.003	-0.015	0.047
	(0.018)	(0.020)	(0.048)
Spline 13+ Months	-0.018	-0.019	-0.003
	(0.031)	(0.033)	(0.094)
Dummy 12th Month	0.148	0.180	0.021
	(0.263)	(0.283)	(0.726)
Dummy 24th Month	0.147	0.389	-0.065
	(0.725)	(0.749)	(0.641)
Non-exempt work-eligible households in the post-DRA period interacted with			
Spline 1-4 Months	-0.017	-0.037 **	0.070 **
	(0.016)	(0.018)	(0.035)
Spline 5-12 Months	0.003	0.009	-0.021
	(0.020)	(0.022)	(0.051)
Spline 13+ Months	0.013	0.023	-0.044
	(0.032)	(0.034)	(0.095)
Dummy 12th Month	-0.101	-0.019	-0.547
	(0.278)	(0.301)	(0.759)
Dummy 24th Month	0.127	-0.093	
	(0.754)	(0.780)	
Time Trend	0.001	0.000	0.001
	(0.001)	(0.001)	(0.001)
Log Likelihood	-92704.33	-73345.10	-19381.31
Number of Monthly Observations	192,814	155,555	37,259

Note: Standard errors are in parentheses. The model for non-blacks omits the dummy variable for month 24 duration for non-exempt work-eligible households in the post-DRA period due to a lack of exits of this group in the 24th month. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 12. Competing Risks Estimates for the Full Sample

	Employment Exits	Other Income Exits	Administrative Exits
Female	-2.274 ***	-2.820 ***	-1.848 ***
	(0.082)	(0.124)	(0.072)
High School Graduate	0.586 ***	0.769 ***	-0.079 ***
	(0.028)	(0.059)	(0.020)
Black	-0.485 ***	-0.390 ***	-0.457 ***
	(0.032)	(0.061)	(0.026)
Age of Household Head	-0.019 ***	-0.023 ***	-0.023 ***
	(0.002)	(0.004)	(0.002)
Age of Oldest Child	0.094 ***	0.163 ***	0.063 ***
	(0.010)	(0.020)	(0.008)
Age of Youngest Child	-0.109 ***	-0.158 ***	-0.035 ***
	(0.010)	(0.018)	(0.007)
Number of Children Age 0-2	-0.512 ***	-1.207 ***	-0.329 ***
	(0.033)	(0.068)	(0.025)
Number of Children Age 3-5	-0.341 ***	-0.798 ***	-0.236 ***
	(0.026)	(0.056)	(0.021)
Number of Children Age 6-11	-0.275 ***	-0.585 ***	-0.104 ***
	(0.035)	(0.072)	(0.026)
Number of Children Age 12-14	-0.300 ***	-0.517 ***	-0.018
	(0.090)	(0.157)	(0.062)
Number of Children Age 15-17	-0.555 ***	-0.681 **	-0.125
	(0.192)	(0.288)	(0.139)
Total Months of Benefit Receipt	0.008 ***	0.003	0.036 ***
	(0.002)	(0.005)	(0.002)
Unemployment Rate	-14.532 ***	2.361 ***	-4.351 ***
	(0.505)	(0.852)	(0.376)
Baseline Duration Dependence			
Spline 1-4 Months	0.318 ***	-0.094 *	0.223 ***
	(0.032)	(0.054)	(0.028)
Spline 5-12 Months	0.056 **	0.114 *	0.072 ***
	(0.027)	(0.061)	(0.020)
Spline 13+ Months	0.086 *	0.073	0.089 ***
	(0.045)	(0.086)	(0.031)
Dummy 12th Month	-0.087	-0.379	0.190
	(0.422)	(1.085)	(0.250)
Dummy 24th Month			1.954 ***
			(0.598)

Non-exempt work-eligible households interacted with			
Spline 1-4 Months	0.164 ***	0.172 ***	0.175 ***
	(0.019)	(0.042)	(0.015)
Spline 5-12 Months	-0.029	-0.096	-0.034 *
	(0.026)	(0.063)	(0.019)
Spline 13+ Months	-0.018	0.002	-0.021
	(0.046)	(0.091)	(0.032)
Dummy 12th Month	0.444	0.592	0.137
	(0.436)	(1.122)	(0.260)
Dummy 24th Month			0.102
			(0.617)
Post-DRA period interacted with			
Spline 1-4 Months	0.060 **	0.077	0.053 ***
	(0.026)	(0.059)	(0.020)
Spline 5-12 Months	0.017	0.023	-0.003
	(0.031)	(0.081)	(0.023)
Spline 13+ Months	-0.048	-0.147	-0.022
	(0.059)	(0.172)	(0.035)
Dummy 12th Month	0.251	-0.633	0.125
	(0.496)	(1.520)	(0.306)
Dummy 24th Month			0.909
			(0.731)
Non-exempt work-eligible households in the post-DRA period interacted with			
Spline 1-4 Months	-0.056 **	0.035	-0.003
	(0.025)	(0.059)	(0.020)
Spline 5-12 Months	0.015	-0.013	-0.010
	(0.034)	(0.085)	(0.024)
Spline 13+ Months	0.022	0.104	0.028
	(0.062)	(0.179)	(0.036)
Dummy 12th Month	-0.203	0.002	-0.046
	(0.521)	(1.606)	(0.324)
Dummy 24th Month			-0.714
			(0.761)
Time Trend	0.003 ***	-0.009 ***	< 0.000
	(0.001)	(0.002)	(0.001)

Note: The omitted category is "no exit." Standard errors are in parentheses. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 13. Competing Risks Estimates for the Black Subsample

	Employment Exits	Other Income Exits	Administrative Exits
Female	-2.913 ***	-3.364 ***	-2.438 ***
	(0.109)	(0.160)	(0.099)
High School Graduate	0.632 ***	0.801 ***	-0.088 ***
	(0.033)	(0.070)	(0.024)
Age of Household Head	-0.014 ***	-0.014 ***	-0.020 ***
	(0.003)	(0.005)	(0.002)
Age of Oldest Child	0.094 ***	0.150 ***	0.064 ***
	(0.011)	(0.022)	(0.009)
Age of Youngest Child	-0.118 ***	-0.185 ***	-0.041 ***
	(0.011)	(0.021)	(0.009)
Number of Children Age 0-2	-0.546 ***	-1.316 ***	-0.356 ***
	(0.038)	(0.078)	(0.029)
Number of Children Age 3-5	-0.347 ***	-0.803 ***	-0.222 ***
	(0.030)	(0.064)	(0.024)
Number of Children Age 6-11	-0.255 ***	-0.504 ***	-0.089 ***
	(0.040)	(0.080)	(0.030)
Number of Children Age 12-14	-0.206 **	-0.403 **	-0.018
	(0.103)	(0.180)	(0.074)
Number of Children Age 15-17	-0.562 ***	-0.637 **	-0.439 **
	(0.206)	(0.319)	(0.177)
Total Months of Benefit Receipt	0.007 ***	0.003	0.038 ***
	(0.003)	(0.005)	(0.002)
Unemployment Rate	-14.771 ***	1.874 **	-4.795 ***
	(0.576)	(0.955)	(0.438)
Baseline Duration Dependence			
Spline 1-4 Months	0.341 ***	-0.033	0.232 ***
	(0.037)	(0.060)	(0.032)
Spline 5-12 Months	0.086 ***	0.068	0.097 ***
	(0.029)	(0.068)	(0.022)
Spline 13+ Months	0.087 *	0.145	0.087 ***
	(0.046)	(0.089)	(0.034)
Dummy 12th Month	-0.061		0.186
	(0.428)		(0.271)
Dummy 24th Month			1.817 ***
			(0.623)
Non-exempt work-eligible households interacted with			
Spline 1-4 Months	0.164 ***	0.164 ***	0.198 ***

	(0.022)	(0.046)	(0.018)
Spline 5-12 Months	-0.039	-0.045	-0.045 **
	(0.029)	(0.069)	(0.021)
Spline 13+ Months	-0.023	-0.050	-0.018
	(0.048)	(0.094)	(0.034)
Dummy 12th Month	0.364		0.076
	(0.445)		(0.283)
Dummy 24th Month			0.232
			(0.644)
Post-DRA period interacted with			
Spline 1-4 Months	0.072 **	0.062	0.084 ***
	(0.029)	(0.065)	(0.023)
Spline 5-12 Months	0.002	0.081	-0.016
	(0.034)	(0.086)	(0.025)
Spline 13+ Months	-0.060	-0.184	-0.023
	(0.062)	(0.173)	(0.037)
Dummy 12th Month	0.200		0.145
	(0.512)		(0.330)
Dummy 24th Month			1.190
			(0.767)
Non-exempt work-eligible households in the post-DRA period interacted with			
Spline 1-4 Months	-0.063 **	0.060	-0.037
	(0.029)	(0.065)	(0.023)
Spline 5-12 Months	0.027	-0.085	-0.004
	(0.036)	(0.091)	(0.027)
Spline 13+ Months	0.039	0.140	0.042
	(0.065)	(0.181)	(0.039)
Dummy 12th Month	-0.041		0.061
	(0.542)		(0.352)
Dummy 24th Month			-1.025
			(0.801)
Time Trend	0.002 **	-0.010 ***	-0.001
	(0.001)	(0.002)	(0.001)

Note: The omitted category is "no exit." Standard errors are in parentheses. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 14. Competing Risks Estimates for the Non-Black Subsample

	Employment Exits	Other Income Exits	Administrative Exits
Female	-1.685 ***	-2.372 ***	-1.335 ***
	(0.140)	(0.203)	(0.120)
High School Graduate	0.397 ***	0.552 ***	-0.097 **
	(0.053)	(0.112)	(0.040)
Age of Household Head	-0.038 ***	-0.051 ***	-0.037 ***
	(0.005)	(0.008)	(0.004)
Age of Oldest Child	0.111 ***	0.229 ***	0.068 ***
	(0.022)	(0.045)	(0.017)
Age of Youngest Child	-0.094 ***	-0.083 **	-0.028 *
	(0.021)	(0.040)	(0.015)
Number of Children Age 0-2	-0.488 ***	-0.901 ***	-0.335 ***
	(0.072)	(0.152)	(0.055)
Number of Children Age 3-5	-0.392 ***	-0.836 ***	-0.359 ***
	(0.059)	(0.128)	(0.047)
Number of Children Age 6-11	-0.395 ***	-0.934 ***	-0.183 ***
	(0.081)	(0.163)	(0.057)
Number of Children Age 12-14	-0.559 ***	-0.975 ***	-0.005
	(0.198)	(0.340)	(0.125)
Number of Children Age 15-17	-1.086 **	-1.290 *	0.389
	(0.539)	(0.707)	(0.257)
Total Months of Benefit Receipt	0.010 *	-0.006	0.027 ***
	(0.006)	(0.011)	(0.004)
Unemployment Rate	-15.010 ***	3.115	-3.829 ***
	(1.104)	(1.941)	(0.779)
Baseline Duration Dependence			
Spline 1-4 Months	0.275 ***	-0.267 **	0.240 ***
	(0.071)	(0.128)	(0.060)
Spline 5-12 Months	-0.108	0.268 **	0.009
	(0.082)	(0.117)	(0.046)
Spline 13+ Months	-0.051		0.150 *
	(0.555)		(0.085)
Dummy 12th Month			0.190
			(0.671)
Dummy 24th Month			2.444 ***
			(0.315)
Non-exempt work-eligible households interacted with			
Spline 1-4 Months	0.166 ***	0.227 **	0.109 ***

	(0.044)	(0.108)	(0.031)
Spline 5-12 Months	0.112	-0.215 *	0.013
	(0.081)	(0.119)	(0.044)
Spline 13+ Months	0.140		-0.072
	(0.555)		(0.085)
Dummy 12th Month			0.306
			(0.686)
Post-DRA period interacted with			
Spline 1-4 Months	0.017	0.203	-0.068
	(0.060)	(0.147)	(0.045)
Spline 5-12 Months	0.159 *	-0.334 *	0.055
	(0.094)	(0.197)	(0.057)
Spline 13+ Months	0.184		-0.028
	(0.561)		(0.104)
Dummy 12th Month			0.036
			(0.840)
Non-exempt work-eligible households in the post-DRA period interacted with			
Spline 1-4 Months	-0.010	-0.122	0.132 ***
	(0.059)	(0.146)	(0.044)
Spline 5-12 Months	-0.140	0.368 *	-0.031
	(0.097)	(0.205)	(0.059)
Spline 13+ Months	-0.204		-0.028
	(0.563)		(0.105)
Dummy 12th Month			-0.459
			(0.877)
Time Trend	0.003 *	-0.008 **	< 0.000
	(0.002)	(0.004)	(0.001)

Note: The omitted category is "no exit." Standard errors are in parentheses. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 15. Simulation Results for All Exits

	Median Spell Length (months)	Fraction of Spells Lasting At Least:		
		Four Months	Six Months	Twelve Months
Full Sample				
Actual Data	4.84	0.64	0.38	0.11
Baseline Simulation	5.01	0.65	0.39	0.14
No DRA Simulation	4.93	0.64	0.38	0.13
All DRA Simulation	5.03	0.65	0.40	0.14
Black Subsample				
Actual Data	4.93	0.66	0.39	0.12
Baseline Simulation	5.14	0.67	0.41	0.14
No DRA Simulation	4.94	0.65	0.38	0.13
All DRA Simulation	5.19	0.67	0.42	0.15
Non-Black Subsample				
Actual Data	4.51	0.59	0.32	0.07
Baseline Simulation	4.61	0.60	0.33	0.09
No DRA Simulation	4.92	0.64	0.38	0.11
All DRA Simulation	4.57	0.59	0.32	0.08

Note: The "No DRA" simulation estimates what would have happened to spell length in the absence of the DRA. The "All DRA" simulation predicts what would have happened had all work-eligible households been subject to the DRA beginning in October 2006. The data used for this table include spells that are right-censored. The statistics from the actual data are calculated using spells that start in the post-DRA period.

Table 16. Simulation Results for Employment Exits

	Median Spell Length (months)	Fraction of Spells Lasting At Least:		
		Four Months	Six Months	Twelve Months
Full Sample				
Actual Data	4.73	0.66	0.34	0.07
Baseline Simulation	4.82	0.64	0.36	0.09
No DRA Simulation	4.74	0.63	0.34	0.08
All DRA Simulation	4.81	0.63	0.39	0.09
Black Subsample				
Actual Data	4.81	0.68	0.36	0.08
Baseline Simulation	4.97	0.66	0.38	0.09
No DRA Simulation	4.77	0.64	0.34	0.07
All DRA Simulation	5.01	0.66	0.39	0.10
Non-Black Subsample				
Actual Data	4.44	0.60	0.27	0.04
Baseline Simulation	4.42	0.57	0.29	0.06
No DRA Simulation	5.14	0.66	0.40	0.10
All DRA Simulation	4.31	0.55	0.27	0.05

Note: The "No DRA" simulation estimates what would have happened to spell length in the absence of the DRA. The "All DRA" simulation predicts what would have happened had all work-eligible households been subject to the DRA beginning in October 2006. Statistics are calculated for spells that end in employment. The data used for this table do not include spells that exit for other reasons or spells that are right-censored. The statistics from the actual data are calculated using spells that start in the post-DRA period.

Table 17. Simulation Results for Other Income Exits

	Median Spell Length (months)	Fraction of Spells Lasting At Least:		
		Four Months	Six Months	Twelve Months
Full Sample				
Actual Data	4.10	0.52	0.25	0.04
Baseline Simulation	3.86	0.47	0.24	0.03
No DRA Simulation	3.70	0.44	0.21	0.03
All DRA Simulation	3.88	0.48	0.25	0.04
Black Subsample				
Actual Data	4.16	0.53	0.26	0.04
Baseline Simulation	3.95	0.49	0.25	0.04
No DRA Simulation	3.92	0.48	0.27	0.05
All DRA Simulation	3.98	0.50	0.24	0.04
Non-Black Subsample				
Actual Data	3.86	0.47	0.20	0.03
Baseline Simulation	3.36	0.36	0.16	0.03
No DRA Simulation	3.35	0.32	0.08	0.00
All DRA Simulation	3.60	0.42	0.24	0.07

Note: The "No DRA" simulation estimates what would have happened to spell length in the absence of the DRA. The "All DRA" simulation predicts what would have happened had all work-eligible households been subject to the DRA beginning in October 2006. Statistics are calculated for spells that end due to increases in other income. The data used for this table do not include spells that exit for other reasons or spells that are right-censored. The statistics from the actual data are calculated using spells that start in the post-DRA period.

Table 18. Simulation Results for Administrative Exits

	Median Spell Length (months)	Fraction of Spells Lasting At Least:		
		Four Months	Six Months	Twelve Months
Full Sample				
Actual Data	4.57	0.59	0.33	0.08
Baseline Simulation	4.54	0.59	0.31	0.08
No DRA Simulation	4.50	0.59	0.31	0.07
All DRA Simulation	4.56	0.59	0.32	0.08
Black Subsample				
Actual Data	4.65	0.60	0.35	0.09
Baseline Simulation	4.67	0.62	0.33	0.08
No DRA Simulation	4.58	0.60	0.31	0.07
All DRA Simulation	4.68	0.62	0.33	0.08
Non-Black Subsample				
Actual Data	4.29	0.55	0.29	0.05
Baseline Simulation	4.30	0.55	0.28	0.06
No DRA Simulation	4.31	0.55	0.29	0.05
All DRA Simulation	4.32	0.55	0.28	0.06

Note: The "No DRA" simulation estimates what would have happened to spell length in the absence of the DRA. The "All DRA" simulation predicts what would have happened had all work-eligible households been subject to the DRA beginning in October 2006. Statistics are calculated for spells that end for administrative reasons. The data used for this table do not include spells that exit for other reasons or spells that are right-censored. The statistics from the actual data are calculated using spells that start in the post-DRA period.

CHAPTER IV

WHAT HAPPENED TO CASH ASSISTANCE FOR NEEDY FAMILIES?

Coauthored with David C. Ribar

1. Goals of Social Assistance

Controversy has surrounded the federally-supported cash assistance program for poor families with children since its inception in the Social Security Act of 1935.[1] Originally called the Aid to Dependent Children (ADC) program, it was rechristened the Aid to Families with Dependent Children (AFDC) program following reforms in 1962 and later the Temporary Assistance for Needy Families (TANF) program following reforms in 1996. Part of the controversy likely stems from the program's costs; empirical studies have found that taxpayer support for transfers falls when the cost of assistance increases (Gramlich 1982; Moffitt 1990; Orr 1976; Ribar and Wilhelm 1999). However, more controversy seems to center on the ways that means-tested programs work at cross-purposes, alleviating the immediate condition of poverty while at the same time encouraging behaviors that can lead families into poverty.

[1] The controversies leading up to the creation of the AFDC program in 1935 are documented in Gordon (1994). Evidence of later controversies can be seen in the major presidential candidates from John Kennedy to George W. Bush calling for major reforms to the program.

What are the goals of cash assistance programs for poor families with children, and given the attendant costs, why would broad sets of taxpayers ever support them? Rational self-interest might be one motivation—taxpayers might want a safety net in place in case they ever fall on hard times. Though reasonable, selfishness seems like an incomplete explanation because few people would ever have the need for this assistance. Hochman and Rodgers (1969) proposed a more universal motivation, theorizing that taxpayers are partially altruistic and care not only about their own well-being but also about the well-being of others, such as disadvantaged families. Even with this explanation, however, the question remains of how best to improve the well-being of disadvantaged families.

On the one hand, transferring money to people improves their well-being by giving them more resources to use to purchase goods and services. On the other hand, transferring money to a family also changes the incentives for that family to earn income on its own and may even encourage the family to expand its needs.

We can consider some of these incentives for work. The vast majority of people get their incomes through work and earnings. If people value the time they spend away from their jobs (or dislike the time they spend at their jobs), transferring money to them will reduce their incentives to work, lowering their earnings and making them poorer in terms of non-transfer income. Worse from an incentives standpoint, the process of means-testing in cash welfare programs causes the eligibility for and amount of assistance to fall as a person's pre-transfer income increases. Means-testing is intended to limit assistance to those who need it most, but it has the unintended effect of acting as an

extra tax on the earnings of program recipients, lowering the rewards associated with work. In some cases, including the TANF programs currently operating in some states today, benefits are docked exactly one dollar for each dollar earned in pre-transfer income, completely eliminating the financial incentive to work among people who command low hourly wages or can only work a few hours per week.

In addition to the short-term effects on work, welfare programs may encourage other behaviors that contribute to poverty over the longer term (Murray 1984). Because the AFDC/TANF program is only available for households with children, it may encourage people to have more children than they otherwise would. Rules that have either limited welfare to single parents or that have made welfare harder to obtain for married parents have the unintended effect of discouraging marriage. The availability of welfare may also reduce the incentives to complete school or to acquire skills. Moreover, a parent's participation in welfare may also influence the future behavior of her children, leading to an intergenerational dependence on assistance. While all of these longer-term effects raise concerns, most empirical research indicates that these effects are modest or negligible (Blank 2002; Moffitt 1992).

As we discuss in this chapter, the U.S. cash welfare system underwent a number of significant reforms to address the programs' deleterious incentives. The reforms were intended to discourage dependency and promote economic self-sufficiency. Most of these reforms, however, took the form of direct or indirect reductions in assistance to families. Examples of direct reductions include imposing life-time limits on the receipt of assistance and ending the entitlement to assistance. Indirect reductions include

conditioning welfare receipt on work or schooling. These restrictive actions raise the question of whether the other goal of welfare programs—to provide help to the disadvantaged—has been compromised.

In the next section of this chapter, we review the reforms that occurred in the AFDC program, starting in the early 1990s. We also discuss the economic circumstances in which these changes occurred. We follow that discussion with a description of trends in outcomes and well-being measures for at-risk families generally, including trends in employment rates, poverty, single parenthood, and welfare participation. Many of these trends indicate that well-being improved on average, at least in the years initially following the reforms. Finally, we examine the conditions and circumstances of the shrinking number of families that continued to rely on cash assistance. Average well-being for these families appears to have suffered in the wake of reform.

2. Policy and economic changes since the early 1990s

AFDC, the cash welfare program in place at the start of the 1990s, was a federal-state partnership. The states operated and administered the assistance programs under a general set of rules and with financial support from the federal government. Within these rules, each state set its own maximum benefit level and determined the maximum level of income that qualified for assistance. States were also responsible for the day-to-day administration of the program. Financial assistance from the federal government took the form of open-ended matching grants in which a dollar of benefit spending by the states was matched by one to as much as three and a half dollars in federal support, depending on the state's relative economic standing. The rules set by the federal government

included the general benefit and eligibility formulas. Following an earlier reform in 1988 (the Family Support Act of 1988), states were required to offer assistance to two-parent families (though under stricter conditions than single-parent families), operate job assistance programs, mandate work among some recipients, offer child care to working recipients, and provide transitional assistance to families who worked their way off the program (Moffitt 1992).

Because of the widespread dissatisfaction with the AFDC program, the federal government began granting states waivers from the program rules, starting in the early 1990s. The waivers were intended to allow states to experiment with different program structures and were generally granted if a state could show that the changes would not cost the federal government more money. Ultimately, 43 states took advantage of this opportunity to reform their own programs and were granted waivers.

The waivers generally included changes in several program elements, with the set of changes being unique to each state. Crouse (1999) categorized the changes into six types: (1) imposition of time limits on the receipt of benefits, (2) changes in the groups covered by mandatory work and training requirements, (3) changes in the amount of time before recipients were required to work, (4) changes in benefit sanctions from not meeting work and program requirements, (5) imposition of "family caps" (families could not collect additional benefits for children born while the family was on assistance), and (6) increases in earnings disregards or decreases in benefit reduction rates. With the exception of the modifications in earnings disregards and benefit formulas, the changes had the effect of reducing the generosity of welfare. Time limits on receipt reduced the

duration of benefits and were intended to shift welfare toward being temporary, rather than permanent assistance. The changes in work rules made welfare harder to obtain by conditioning receipt on employment or other work activities, including job search, education, training, and community work experience, and were intended to promote employment. The changes in earnings disregards and benefit formulas were also intended to promote employment but through the positive incentive of letting welfare recipients keep more of their welfare benefits as their earnings increased.

In 1996, Congress passed the Personal Responsibility and Work Opportunity Reconciliation Act (PRWORA), which replaced the states' AFDC programs with TANF programs and also reformed other assistance programs. The PRWORA changed the cash assistance system in several fundamental respects. First, it restricted federally-supported welfare by limiting life-time receipt to five years and by imposing new work requirements on recipients. Second, it changed the federal support from an unlimited matching grant to a fixed block grant. The block grant would be conditioned on a proportion of the state's caseload meeting work requirements and on the state expending a portion of its own funds. Third, the PRWORA eliminated the entitlement to assistance—once states had exhausted their annual block grants and met a maintenance-of-effort requirement with their own funds, they were no longer required to pay benefits. Fourth, it removed other restrictions from the states' operation of their programs, allowing them to incorporate their own reforms. Thus, the legislation included more restrictions on states in some elements of their assistance programs but more flexibility in other areas.

States had to submit their TANF plans to the Department of Health and Human Services (DHHS) for approval. The first TANF programs were implemented in September 1996, and all of the states had implemented programs by the start of 1998. Where they had the option to do so, some states adopted elements from their AFDC waivers, while others crafted entirely new programs (Crouse 1999).

TANF was initially authorized for five years. When TANF came up for renewal, legislative disagreements over the types of subsequent reforms led to a series of one-year continuations. During this period, the U.S. General Accountability Office (GAO) issued a report identifying inconsistent and suspect practices across states in the types of activities that could count toward meeting the federal work requirements and in the internal controls used to verify the accuracy of reported work hours (U.S. GAO 2005). For example, the GAO found that some states were counting bed rest, exercise, smoking cessation, and massage therapy as "work readiness activities." When TANF was reauthorized as part of the Deficit Reduction Act of 2005 (DRA), work requirements were strengthened through three main channels: (1) by requiring that families in separate state-funded programs meet the work requirements, (2) by increasing the percentage of recipients required to work in each state, and (3) by standardizing work eligibility and the activities that satisfy the work requirement across states. These changes went into effect on October 1, 2006 (71 Federal Register 125 (2006)).

The changes in the AFDC and TANF programs were accompanied by changes in other assistance programs. The Earned Income Tax Credit (EITC), a refundable tax credit available to people with low to moderate levels of earned income, was expanded

significantly in 1993, when AFDC waivers were becoming more common. The EITC significantly reduced the tax burden of low-income households, thus increasing the incentives to work. The primary welfare reform legislation, the PRWORA, also made changes to other assistance programs. For example, the PRWORA altered the Food Stamp Program to reduce benefits, eliminate eligibility for most immigrants, and impose work requirements and time limits on able-bodied adults without dependents. Later legislation restored eligibility for many immigrants and made food stamps easier to obtain. In 1997, Congress passed State Children's Health Insurance Programs (SCHIP), which expanded Medicaid coverage to children of low-income, working parents. Families who received or had recently transitioned off welfare automatically qualified to receive Medicaid benefits, creating yet another incentive to become eligible for welfare (Yelowitz 1995). The expansion of these benefits to low-income, working families was partly intended to improve the incentives for work. A study by Moffitt and Scholz (2009) found that the constellation of reforms had the effect of shifting the distribution of transfers away from the poorest single- and two-parent households and toward near-poor households and households with disabled members.

The changes in cash assistance also occurred in the context of a growing economy. The AFDC waivers began to be implemented as the U.S. was emerging from a recession in 1990-1991, and the PRWORA was enacted just a few years into what would become the longest economic expansion in U.S. history. The national unemployment rate dropped from a peak of 7.8 percent in July 1992 to 5.1 percent at the time PRWORA was enacted in August 1996. The national unemployment rate continued to fall throughout the

late 1990s, reaching a low of 3.8 percent in April 2000. A mild recession in 2001 caused unemployment to rise, but only to a level of 6.3 percent. From the middle of 2005 until the start of the Great Recession at the end of 2007, unemployment remained at or below 5 percent (U.S. Department of Labor 2011). A rising tide might not lift all boats, but the robust job market that characterized most of the 1990s and much of the 2000s undoubtedly made the policy goals of promoting work and economic self-sufficiency among the disadvantaged easier to achieve. Conversely, with unemployment soaring during the Great Recession, the well-being of disadvantaged families has likely deteriorated.

3. Trends in general well-being outcomes

One measure of the general well-being of at-risk families is the welfare caseload, the number of families receiving welfare benefits. As the caseload decreases, we might believe that the well-being of at-risk families has increased. Prior to the start of the AFDC waiver period in fiscal year (FY) 1990, there were approximately 3.9 million households in the U.S. receiving cash welfare payments.[1] The initial waivers were requested and granted at a time of expanding caseloads. However, after reaching a peak in FY 1994, the caseload began a fourteen-year decline. In FY 1995, just one year before PRWORA was enacted, there were roughly 4.8 million households receiving AFDC benefits, and by FY 2008, the caseload had dropped to approximately 1.6 million households.

[1] These and subsequent figures include all 50 states and the District of Columbia but exclude U.S. territories.

Of course, if caseload reduction were the foremost goal of assistance policy, we could achieve that goal overnight by eliminating TANF altogether. Instead, welfare assistance focuses on alleviating poverty. The poverty rate among those at risk of becoming dependent on welfare, single mothers with dependent children, has exhibited clear trends over the past twenty years. In 1991, over 47 percent of all single mothers with children under 18 were living in poverty. As economic conditions improved over the 1990s, so did the poverty rate among this group of women, and it reached an all-time low of 33 percent in 2000. However, between 2000 and 2008, conditions worsened for single mothers, and the percentage of those living in poverty grew to just over 37 percent in 2008 (U.S. Census Bureau 2011b).

The employment rate of single mothers is a measure of well-being closely tied to poverty and self-sufficiency. In particular, the earned income from employment may be enough to elevate a woman and her family out of poverty and into self-sufficiency so that she is no longer dependent on welfare assistance. The employment of single mothers improved throughout the 1990s. In 1990, almost 56 percent of single mothers in the U.S. were working, and by 2000, this figure had reached almost 63 percent. However, by 2003, the employment rate for this group had dipped to approximately 60 percent, and it remained between 60 and 61 percent from 2003 to 2008.[2]

The welfare program primarily benefits single-parent families. Thus, a key measure of well-being that puts a woman at risk of becoming dependent on assistance is

[2] The employment figures were calculated from the Integrated Public Use Microdata Series of the U.S. Current Population Survey (see King et al. 2010).

female headship or the birth of a child out of wedlock. Data from the Centers for Disease Control and Prevention (2011) show that nonmarital births have been generally increasing over the past twenty years. In 1990, 28 percent of all births were to unmarried mothers. By 2000, this figure had surpassed 33 percent. The percentage of nonmarital births continued to rise over the next eight years, reaching just under 41 percent by 2008. The fertility rate of unmarried women has also risen since 1990; however, the increase occurred primarily in more recent years. In 1990, there were fewer than 44 live births to unmarried women age 15-44 per 1000 such women in the population. Between 1990 and 2004, the rates fluctuated between 43 and 46; however, after 2004, the fertility rate of unmarried women began a marked increase. By 2008, there were almost 53 live births to unmarried women age 15-44 per 1000 such women in the population.

As the fertility rate of unmarried women increases, so too does the proportion of women at risk of becoming dependent on the welfare system. However, there have been some positive trends. In particular, the fertility rate of unmarried teenage women has been decreasing over the past 20 years. In 1990, there were approximately 30 births to single women age 15-17 per 1000 single women age 15-17 in the population. Between 1990 and 1995, this figure fluctuated between 30 and 32; however, beginning in 1996, it started to decline. In 2005, the fertility rate of single teens had dropped to just under 20 live births per 1000 teen women. Although the fertility rate of single teens did increase slightly between 2005 and 2008, it remained below 21 live births per 1000 teen women during these years (Martin et al. 2010).

Numerous studies have examined the ways in which welfare reform and other policy and economic changes have contributed to these different trends, with the findings from the studies being summarized in several comprehensive reviews by Blank (2002), Grogger and Karoly (2005), and Moffitt (2002). The studies have generally found that welfare reform has played a role in reducing the assistance caseload, increasing employment, and increasing earnings. There is less agreement, however, regarding which specific reform components have been responsible. For example, the Council of Economic Advisors (1997) found that only work sanction waivers had a significant impact in reducing welfare recipiency. However, Moffitt (1999) concluded that work requirements and family caps were important, while Grogger and Karoly (2005) concluded that mandatory work requirements and financial incentives were crucial. In contrast to the research on caseload and economic outcomes associated with welfare reform, the findings regarding the associations with demographic outcomes, such as marriage and single-parenthood has been more equivocal.

4. Data on recipients

While the economic well-being of families at risk of becoming dependent on welfare did improve between 1990 and 2008, those actually receiving welfare have not experienced the same progress. In order to show that this is the case, we will make use of administrative data from the DHHS on recipients of AFDC and TANF assistance. The DHHS publishes tables and makes available annual, public-use data on a sample of households from the AFDC and TANF programs in each state (U.S. DHHS 2011a; U.S. DHHS 2011b; U.S. DHHS 2011f). Information on active AFDC households is available

from fiscal year (FY) 1967 through FY 1997, while data on both active and closed TANF cases are available from FY 2000 through FY 2008, to date. Data on closed TANF cases are reported for the households' last month of TANF cash receipt. In this chapter, we will focus on AFDC data from fiscal years 1990 and 1995 and TANF data from fiscal years 2000, 2005, and 2008 for all fifty states and the District of Columbia.

The DHHS data encompass a rich set of variables for analysis. For each household in the sample, the data contain information such as household size, assistance received, and financial resources. The data also contain detailed information on both recipient and non-recipient individuals in each household. The person-level data include relationships within the household, demographic information, education level, employment status, citizenship, and disability receipt. The broad range of both household-level and person-level characteristics available in this dataset make it particularly well suited for this chapter. The data on earned and unearned income, education, and employment allow us to study trends in the economic well-being of welfare recipients, while the extensive data on cash and non-cash assistance allow us to study trends in recipiency over the past twenty years of welfare reform.

5. The rise in child-only cases

One of the most noticeable trends in the cash assistance caseload has been the steady rise in the proportion of cases in which children are the only members of the household receiving assistance—so-called "child-only" cases. The DHHS (2010a) tabulations indicate that in FY 1990, child-only cases comprised just under one-eighth of the welfare caseload. By 1995, just prior to the enactment of the PRWORA, the fraction

had increased to slightly more than one-sixth. By FY 2000, the fraction had reached one-third, and by FY 2008, the fraction had climbed to slightly over one-half. The absolute number of child-only cases peaked in FY 1996 at nearly 1 million, fell in the first few years after the PRWORA, but began climbing in FY 1999. In FY 2008, there were slightly more than 800,000 child-only cases (Charlesworth, Hercik, and Kakuska 2011).

Child-only cases arise under a number of circumstances. In about two-thirds of these cases in FY 2008, a parent is present in the household but is not included in the welfare assistance unit (U.S. DHHS 2010b). A parent would not be included in the assistance unit if she were receiving disability on her own, was an immigrant, had been sanctioned for not complying with a work or other requirement, or was excluded for other reasons. In the remaining one-third of child-only cases, the parent is not present in the household, and the children are being cared for by a relative or other caretaker.

There are several policy and well-being concerns surrounding child-only cases. First and foremost, these cases generally receive fewer benefits than they might otherwise because at least one fewer person is included in the assistance unit. In cases in which the parent is present in the household, this will mean fewer resources for the family. In cases where other caretakers are responsible for the children, the hardships can extend to other family members and may discourage some people from caring for relatives. The second concern is that these are cases where there are few options for promoting self-sufficiency. Third, the generally upward trend in the number of child-only cases since 1999 is an indicator of underlying economic needs.

6. Outcomes for parent-headed cases

We use the administrative micro-data from the DHHS to construct tables describing how assistance, economic outcomes, and other characteristics for more traditional parent-headed cases have changed over time. Statistics for these families are reported in Table 19. The figures indicate that the average amount of monthly cash assistance for these households has declined markedly over time from $662 (in constant 2008 dollars) per family in FY 1990 to $413 in FY 2008, a 28 percent drop. By way of comparison, the poverty threshold for a family of three in 2008 was $1,467, so average benefits fell from 45 percent of the poverty standard in FY 1990 to just 28 percent of the poverty standard in FY 2008.

However, cash assistance was not the only resource that welfare families could draw on. The reforms to the welfare system were intended to increase work among recipients. These efforts were partially successful. While fewer than one in nine welfare cases with an adult recipient had earned income in FY 1990, approximately one quarter of such cases had earned income in FY 2008. While the percentage increase in recipients with earnings was sizeable, the percentage was still well short of the 50 percent goal in the PRWORA and DRA.

The increased work effort among recipients led to higher levels of earnings. Average monthly earnings in adult-headed cash assistance households were $57 in FY 1990 but nearly $200 by FY 2008. Over the same period, average monthly unearned income was little changed—$37 in FY 1990 and $30 in FY 2008. Because of the increase in earnings, total monthly pre-transfer incomes rose from an average of $94 in FY 1990

to $229 in FY 2008. These increases in pre-transfer income offset some but not all of the decrease in assistance benefits. Combining earned income, unearned income, and welfare assistance amounts, total monthly post-transfer incomes decreased from $776 (52 percent of the poverty threshold) in FY 1990 to $642 (44 percent of the poverty threshold) in FY 2008. Thus TANF families were working harder after welfare reform but receiving less money overall.

The picture does not change much when other types of assistance are taken into account. The percentage of adult-headed welfare families that also received food stamps fluctuated only a little over the period. However, the value of the food stamps received increased slightly because of the fall in total post-transfer incomes (the means-testing formula for food stamp benefits includes the value of cash assistance) and because of the shift in composition of income toward earnings (the food stamp formula provides an extra deduction for earned income). Welfare families were generally categorically eligible for Medicaid before and after the reform, and thus we do not see much change in the receipt of medical assistance over this period. Conversely, we do observe a decline in the proportion of welfare families receiving housing assistance. Finally, welfare families with earnings benefited from more generous payments from the EITC. For instance, a single-parent household with two children and monthly earnings of $200 (the average earnings for TANF families in 2008) might have qualified for an EITC worth the equivalent of $80 per month in 2008. In 1990, the EITC for the same-size family with $57 in monthly earnings (the average in 11990) would have been just $8.

When we examine the demographic characteristics of adult-headed welfare households, we see that the average number of people in each household shrank from 3.7 in FY 1990 to 3.0 in FY 2008. However, the average number of members of each assistance unit was essentially unchanged at about three in each period. The number of household members would be higher than the number of case members if older relatives or adult children were living with the recipient family.

The average age of household heads also changed little over the period, with the average age being around 30 years old. The proportion of adult heads who were white or black decreased slightly from FY 1990 to FY 2008, while the proportion that was Hispanic increased slightly.

One surprising trend is that among household heads who are welfare recipients, the percentage who have completed high school or an equivalent credential doubled from 27 percent in FY 1990 to 57 percent in FY 2008. The increase was much steeper than the overall rise in educational attainment in the U.S. (figures from the Census Bureau (2011c) indicate that the proportion of women aged 25 years or older who had completed high school rose from 77 percent in 1990 to 87 percent in 2008). We might expect the caseload to have become less skilled over time, as adults with more skills and greater earnings potential left or avoided welfare to pursue other opportunities. However, the statistics indicate that the remaining caseload actually became more highly educated, which suggests that less-skilled adults were disproportionately dropped or diverted from the program. Reduced participation among the least-skilled is consistent with evidence from the Food Stamp Program that exceptionally disadvantaged households have more

trouble completing applications, keeping up with program paperwork, and complying with other participation requirements (Ribar and Edelhoch 2008; Ribar and Swann 2011). The unsettling implication of these trends is that cash assistance might be bypassing the most disadvantaged and truly needy households.

Additional evidence that selective attrition from the caseload occurred from both the high and low ends of the skill distribution comes from examining the characteristics of closed cases (welfare leavers) in the month before they exited. Characteristics of these cases are reported in Table 20. The statistics indicate that cases that were about to close had higher levels of work effort and higher levels of earnings than cases generally. In FY 2008, nearly a third of adult-headed cases that were about to close had earnings, while only a quarter of the general adult caseload had earnings. Average earnings in cases that were about to close were $374 compared to $199 in the general caseload. Though work and earnings for closing cases were higher in a relative sense, it is important to note that two-thirds of the adult cases closed without any earnings. At the same time, education levels for leavers in FY 2008 were lower than education levels for the general caseload, which may be indicative of selection among less-skilled individuals. Also, the incidence and amount of earnings among cases that were closing were each lower in FY 2008 than in FY 2000.

We can also examine the reasons why cases were closed. The DHHS (2011c) reports that in FY 2008, 15 percent of cases closed because they failed to cooperate with program rules; 13 percent closed because of sanctions; and 2 percent closed because of state or federal time limits. Only 20 percent of cases closed because of increased

earnings. Thus the majority of cases closed not because the families achieved self-sufficiency but rather because they failed to meet program requirements.

7. Conclusions

Concerns that cash assistance was fostering dependency and eroding responsibility among recipient families led the states and ultimately the federal government to undertake fundamental reforms of the welfare system in the 1990s. The reforms were intended to promote economic self-sufficiency by requiring recipients to work and by placing time limits on assistance. The work rules and some other program requirements were backed up by the possible loss of benefits for those who failed to comply. More generally, disadvantaged families lost their entitlement to assistance. Although some of the reforms adopted by some states included positive incentives for work, such as higher earnings disregards and lower benefit reduction rates in calculating assistance, most of the reforms were punitive, involving a loss of benefits if certain behaviors did not occur. The negative nature of these reforms created countervailing concerns that the assistance function of the cash welfare system would be compromised and that the well-being of many disadvantaged families would suffer.

The evidence presented in this chapter indicates that many disadvantaged families experienced beneficial outcomes following the implementation of the reforms. The most notable beneficial effects were a sharp decrease in the proportion of single-parent households in poverty from just under half prior to the reforms to a third by 2000 and an increase in employment from 54 percent of single mothers prior to the reforms to 63 percent in 2000. Since 2000, poverty has crept back up, while employment has decreased.

Even more dramatic has been the sharp fall in the cash assistance caseload from 4.8 million families in 1995 to 1.6 million families in 2008, a two-thirds decrease. Research indicates that the reforms contributed substantially to these outcomes.

However, the reforms have also been associated with deleterious effects. The fall in the caseload has been accompanied by profound changes in its composition. Child-only cases have increased from a small fraction of the caseload prior to welfare reform to half of the caseload in 2008. In most of the current cases where household children qualify for assistance but adults do not, parents have been sanctioned off the program or have been made ineligible for benefits. Thus, these families are receiving fewer benefits than they would have prior to the reforms. About a third of the child-only cases, however, represent children who are living with a caregiver other than their parents. In either circumstance, child-only cases represent a growing policy challenge because there are few ways for the welfare system to incentivize the economic self-sufficiency for the family.

Among the more traditional, parent-headed assistance cases, work and earnings have increased, though the incidence of work and the average amount of earnings remain very modest. In 2008, one out of four parent-headed assistance cases received earnings, and the average monthly earnings among those who worked was about $800, far below the poverty threshold for these families. At the same time, average benefits have fallen faster than earnings have grown. The net result is that parent-headed assistance families are working harder than they were before the reforms were enacted but receiving smaller amounts of post-transfer income.

Evidence also suggests that the reductions in the caseload may have occurred disproportionately among households with the least skills. The average level of education among adult heads of assistance cases, while still lower than the population average, has risen substantially since the enactment of welfare reform. Only about 20 percent of families who leave welfare have done so because earnings made them ineligible; a much larger fraction leaves because of sanctions and other program rules. Indeed, two-thirds of families do not have any earnings in the month before they leave the program.

The picture that emerges is of a program that has all but abandoned its assistance mission. The Census Bureau (2011a) estimates that 23.7 million people (including 13.5 million children) were members of families with children that had incomes below the poverty line in 2008. Of these, the TANF program provided cash benefits to 3.7 million people, or fewer than one in six. The cash benefits that were provided were only a fraction of the income needed to reach the poverty level. Food, energy, housing, and medical in-kind assistance programs still assist substantial proportions of disadvantaged families, but the TANF program does not.

Table 19. Characteristics of Active AFDC/TANF Cases with Adult Recipients

	FY 1990	FY 1995	FY 2000	FY 2005	FY 2008
Household Composition					
Number of household members	3.7	3.8	3.2	3.0	3.0
Number of recipients in the household	3.1	3.1	3.1	2.9	2.9
Percent with 2 or more adult recipients	7.3%	8.5%	6.0%	4.5%	7.3%
Percent headed by teen parents		0.2%	0.5%	0.4%	0.3%
Percent with recipient adults receiving disability benefits	0.9%	1.4%	0.9%	0.7%	0.8%
Percent with recipient children receiving disability benefits	0.1%	1.5%	0.5%	0.6%	0.9%
Household Public Assistance Receipt					
Amount of cash benefits received	$662	$556	$458	$431	$413
Months of cash benefit receipt			23.1	25.6	26.0
Percent receiving food stamps	91.2%	94.4%	89.2%	91.1%	89.4%
Percent receiving medical assistance			99.4%	97.7%	97.5%
Percent living in public housing	9.9%	8.3%	7.1%	5.9%	5.0%
Percent receiving a rent subsidy	15.4%	15.9%	12.7%	13.2%	10.3%
Household Income and Resources					
Percent with earned income	10.6%	13.4%	25.2%	20.5%	24.1%
Earned income amount [a]	$57	$83	$212	$156	$199
Percent with unearned income	9.6%	13.7%	9.2%	8.5%	9.5%
Unearned income amount [b]	$37	$58	$35	$33	$30
Percent with any income	19.1%	25.5%	32.3%	27.4%	31.5%
Total income amount	$94	$141	$248	$189	$229
Percent with cash resources/liquid assets	14.4%	15.2%	13.0%	13.2%	11.9%
Cash resources/Liquid assets	$33	$36	$43	$28	$22
Characteristics of Heads of Households					
Age	29.9	30.7	31.0	30.4	29.9
White [c]	40.1%	38.1%	32.3%	36.3%	35.8%
Black [c]	40.0%	37.0%	39.2%	40.1%	37.2%
Hispanic [c]	14.5%	18.3%	21.5%	18.8%	21.9%
Other/missing race [c]	4.1%	4.6%	5.9%	4.4%	5.0%
Completed high school	26.8%	34.5%	51.2%	59.3%	56.8%
US citizen	92.9%	87.8%	91.0%	93.4%	93.1%
Number of households (millions)	3.4	3.9	1.5	1.0	0.8

Note: The statistics are authors' calculations from DHHS administrative data from the AFDC and TANF programs. All dollar figures are expressed in constant 2008 amounts deflated using the CPI-U.

[a] In FY 1990-1995, earned income is computed for all recipients age 18 and older; in FY 2000-2008, earnings of 18 year olds enrolled full-time in secondary school are excluded.

[b] In FY 1990-1995, unearned income does not include housing assistance; in FY 2000-2008, it does.

[c] The data on race and ethnicity are mutually exclusive in FY 1990-1995 but not in FY 2000-2008. For comparability, race was assigned in FY 2000-2008 as Hispanic, non-Hispanic white, non-Hispanic black, and non-Hispanic other.

Table 20. Characteristics of Closed TANF Cases with Adult Recipients

	FY 2000	FY 2005	FY 2008
Household Composition			
Number of household members	3.0	2.9	3.0
Number of recipients in the household	2.8	2.8	2.8
Number of recipient children in the household	1.7	1.7	1.7
Household Public Assistance Receipt			
Percent receiving food stamps	69.6%	82.3%	83.5%
Percent receiving medical assistance	94.3%	93.6%	93.6%
Percent living in public housing	9.2%	5.1%	4.8%
Percent receiving a rent subsidy	8.4%	10.1%	9.7%
Percent receiving subsidized child care	17.2%	10.7%	10.4%
Household Income and Resources			
Percent with earned income	39.5%	32.4%	32.0%
Earned income amount [a]	$465	$362	$374
Percent with unearned income	17.3%	19.6%	20.7%
Unearned income amount [b]	$93	$106	$112
Percent with any income	51.2%	46.6%	46.5%
Total income amount	$558	$468	$486
Characteristics of Heads of Households			
Age	30.6	30.3	30.7
White [c]	37.4%	38.6%	38.1%
Black [c]	37.0%	35.0%	34.9%
Hispanic [c]	17.6%	20.5%	21.7%
Other/missing race [c]	7.9%	5.8%	5.4%
Married	24.8%	23.3%	20.8%
Completed high school	51.4%	59.6%	50.8%
US citizen	91.9%	93.7%	93.3%
Number of households (millions)	1.5	1.4	1.3

Note: The statistics are authors' calculations from DHHS administrative data from the TANF program. All dollar figures are expressed in constant 2008 amounts deflated using the CPI-U.

[a] Earnings of 18 year olds enrolled full-time in secondary school are excluded from household earned income.

[b] Unearned income includes housing assistance.

[c] The data on race and ethnicity are not mutually exclusive. Mutually exclusive race was assigned as Hispanic, non-Hispanic white, non-Hispanic black, and non-Hispanic other.

REFERENCES

Allison, Paul D. 1984. *Event History Analysis: Regression for Longitudinal Event Data*. Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 07-046. Beverly Hills: Sage Publications.

Balkrishnan, R., M. A. Hall, D. Mehrabi, J. Chen, S. R. Feldman, and A. B. Fleischer. 2002. "Capitation Payment, Length of Visit, and Preventive Services: Evidence from a National Sample of Outpatient Physicians." *American Journal of Managed Care*, 8(4): 332-340.

Beaulieu, N. D. and D. R. Horrigan. 2005. "Putting Smart Money to Work for Quality Improvement." *Health Services Research*, 40(5): 1318-1334.

Blank, Rebecca M. 2002. "Evaluating Welfare Reform in the United States." *Journal of Economic Literature*, 40(4): 1105-1166.

Blank, Rebecca M. and Patricia Ruggles. 1996. "When Do Women Use Aid to Families with Dependent Children and Food Stamps? The Dynamics of Eligibility Versus Participation." *The Journal of Human Resources*, 31(1): 57-89.

Bloom, Dan and Charles Michalopoulos. 2001. "How Welfare and Work Policies Affect Employment and Income: A Synthesis of Research." New York: Manpower Demonstration Research Corporation.

Campbell, S., D. Reeves, E. Kontopantelis, and E. Middleton. 2007. "Quality of Care in England with the Introduction of Pay for Performance." *The New England Journal of Medicine*, 357(2): 181.

Centers for Disease Control and Prevention, National Center for Health Statistics. 2011. VitalStats. Accessed June 9. http://www.cdc.gov/nchs/vitalstats.htm.

Charlesworth, Leanne, Jeanette Hercik, and Courtney Kakuska. 2011. "TANF Child-Only Cases Trends and Issues." Accessed June 9. http://peerta.acf.hhs.gov/pdf/child_only.pdf.

Council of Economic Advisors. 1997. "Explaining the Decline in Welfare Recieipt, 1993-1996." Washington, D.C.: Council of Economic Advisors.

Crouse, Gil. 1999. "State Implementation of Major Changes to Welfare Policies, 1992-1998." http://aspe.hhs.gov/hsp/waiver-policies99/policy_cea.htm.

Devlin, R.A. and S. Sarma. 2008. "Do Physician Remuneration Schemes Matter? The Case of Canadian Family Physicians." *Journal of Health Economics*, 27(5): 1168-1181.

Dumont, E., B. Fortin, N. Jacquemet, and B. Shearer. 2008. "Physicians' Multitasking and Incentives: Empirical Evidence from a Natural Experiment." *Journal of Health Economics*, 27(6): 1436-1450.

Dyke, Andrew et al. 2006. "The Effects of Welfare-to-Work Program Activities on Labor Market Outcomes." *Journal of Labor Economics*, 24(3): 567-607.

Elliott, V.A. 2012. "3 Steps to Quality Pay for Physicians." *American Medical News*. American Medical Association, posted July 23. www.ama-assn.org/amednews/2012/07/23/bisa0723.htm

Fairbrother, G., K. L. Hanson, S. Friedman, and G. C. Butts. 1999. "The Impact of Physician Bonuses, Enhanced Fees, and Feedback on Childhood Immunization Coverage Rates." *American Journal of Public Health*, 89(2): 171-175.

Fairbrother, G., M. J. Siegel, S. Friedman, P. D. Kory, and G. C. Butts. 2001. "The Impact of Financial Incentives on Documented Immunization Rates in the Inner City: Results of a Randomized Controlled Trail." *Ambulatory Pediatrics*, 1 (4): 206-212.

Fang, Hanming and Michael P. Keane. 2004. "Assessing the Impact of Welfare Reform on Single Mothers." *Brookings Papers on Economic Activity*, 2004(1): 1-95.

Frogner, Bianca, Robert Moffitt, and David C. Ribar. 2009. "How Families Are Doing Nine Years after Welfare Reform. 2005 Evidence from the Three-City Study." In James P. Ziliak (ed) *Welfare Reform and Its Long-Term Consequences for America's Poor*. Cambridge University Press.

Gaynor, M. and P. Gertler. 1995. "Moral Hazard and Risk Spreading in Partnerships." *RAND Journal of Economics*, 26(4): 591-613.

Gaynor, M. and M. Pauly. 1990. "Compensation and Productive Efficiency in Partnerships: Evidence from Medical Group Practices." *The Journal of Political Economy*, 98(3): 544-573.

Glied, S. and J. G. Zivin. 2002. "How Do Doctors Behave When Some (but not all) of their Patients are in Managed Care?" *Journal of Health Economics*, 21(2): 337-353.

Gordon, Linda. 1994. *Pitied But Not Entitled: Single Mothers and the History of Welfare*. New York: Free Press.

Grady, K. E., J. P. Lemkau, N. R. Lee, and C. Caddell. 1997. "Enhancing Mammography Referral in Primary Care." *Preventive Medicine*, 26(6): 791-800.

Gramlich, Edward. 1982. "An Econometric Examination of the New Federalism." *Brookings Papers on Economic Activity* 1982 (2): 327-60.

Grogger, Jeffrey and Lynn Karoly. 2005. *Welfare Reform: Effects of a Decade of Change.* Cambridge, MA: Harvard University Press.

Grumbach, K., D. Osmond, K. Vranizan, D. Jaffe, and A. Bindman. 1998. "Primary Care Physicians' Experience of Financial Incentives in Managed-Care Systems." *The New England Journal of Medicine*, 339(21): 1516-1521.

Grytten, J., D. Holst, and I. Skau. 2009. "Incentives and Remuneration Systems in Dental Services." *International Journal of Health Care Finance and Economics*, 9(3): 259-278.

Harris, Kathleen Mullan. 1993. "Work and Welfare Among Single Mothers in Poverty." *The American Journal of Sociology*, 99(2): 317-352.

Heckman, J. and B. Singer. 1984. "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data." *Econometrica*, 52(2): 271-320.

Hochman, Harold and James Rodgers. 1969. "Pareto Optimal Redistribution." *American Economic Review* 59 (4): 542-57.

King, Miriam, Steven Ruggles, J. Trent Alexander, Sarah Flood, Katie Genadek, Matthew B. Schroeder, Brandon Trampe, and Rebecca Vick. 2010. *Integrated Public Use Microdata Series, Current Population Survey: Version 3.0*. [Machine-readable database], Minneapolis: University of Minnesota. http://cps.ipums.org/cps/index.shtml.

Klerman, Jacob Alex and Arleen Leibowitz. 1994. "The Work-Employment Distinction among New Mothers." *The Journal of Human Resources*, 29(2): 277-303.

Lazear, E. 2000. "Performane Pay and Productivity." *The American Economic Review*, 90(5): 1346-1361.

Lillard, L. A. and C. W. A. Panis. 2003. *aML Multilevel Multiprocess Statistical Software, Version 2.0*. Los Angeles, California: EconWare.

Martin, Joyce A., Brady E. Hamilton, Paul D. Sutton, Stephanie J. Ventura, T.J. Matthews, and Michelle J.K. Osterman. 2010. "Births: Final Data for 2008." National Vital Statistics Reports 59 (1), Hyattsville, MD: National Center for Health Statistics.

Melichar, L. 2009. "The Effect of Reimbursement on Medical Decision Making: Do Physicians Alter Treatment in Response to a Managed Care Incentive?" *Journal of Health Economics*, 28(4): 902-907

Moffitt, Robert. 1990. "Has State Redistribution Policy Grown More Conservative?" *National Tax Journal* 43 (2): 123-42.

Moffitt, Robert. 1992. "Incentive Effects of the U.S. Welfare System: A Review." *Journal of Economic Literature* 30 (1): 1-61.

Moffitt, Robert A. 1996. "The Effect of Employment and Training Programs on Entry and Exit from the Welfare Caseload." *Journal of Policy Analysis and Management*, 15(1): 32-50.

Moffitt, Robert. 1999. "The Effect of Pre-PRWORA Waivers on AFDC Caseloads and Female Earnings, Income, and Labor Force Behavior." In *Economic Conditions and Welfare Reform*, edited by Sheldon H. Danziger, 91-118. Kalamazoo, MI: Upjohn Institute.

Moffitt, Robert. 2002. "The Temporary Assistance for Needy Families Program." NBER Working Paper no. w8749, National Bureau of Economic Research, Cambridge, MA.

Moffitt, Robert A. 2003. "The Temporary Assistance for Needy Families Program." in Robert A. Moffitt (ed) *Means-Tested Transfer Programs in the United States*. University of Chicago Press (pp. 291-364).

Moffitt, Robert and John Karl Scholz. 2009. "Trends in the Level and Distribution of Income Support." NBER Working Paper no. w15488, National Bureau of Economic Research, Cambridge, MA.

Mroz, T. 1999. "Discrete Factor Approximations in Simultaneous Equation Models: Estimating the Impact of a Dummy Endogenous Variable on a Continuous Outcome." *Journal of Econometrics*, 92(2): 233-274.

Mullen, K. J., R. G. Frank, and M. B. Rosenthal. 2010. "Can You Get What You Pay For? Pay-for-Performance and the Quality of Healthcare Providers." *RAND Journal of Economics*, 41(1): 64-91.

Murray, Charles. 1984. *Losing Ground: American Social Policy, 1950-1980.* New York: Basic Books.

Orr, Larry. 1976. "Income Transfers as a Public Good: An Application to AFDC." *American Economic Review* 66 (3): 359-71.

Pavetti, LaDonna, Linda Rosenberg, and Michelle Derr. 2009. "Understanding Temporary Assistance for Needy Families Caseloads After Passage of the Deficit Reduction Act of 2005." Washington, D.C.: Mathematica Policy Research, Inc.

Petersen, L. A., L. D. Woodard, T. Urech, C. Daw, and S. Sookanan. 2006. "Does Pay-for-Performance Improve the Quality of Health Care?" *Annals of Internal Medicine*, 145(4): 265-272.

Prendergast, C. 1999. "The Provision of Incentives in Firms." *Journal of Economic Literature*, 37(1): 7-63.

"Reauthorization of the Temporary Assistance for Needy Families Program: Interim Final Rule," 71 Federal Register 125 (29 June 2006), pp. 37454 - 37483.

Ribar, David and Marilyn Edelhoch. 2008. "Earnings Volatility and the Reasons for Leaving the Food Stamp Program." In *Income Volatility and Food Assistance in the United States*, edited by Dean Jolliffe and James Ziliak, 63-102. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Ribar, David and Christopher Swann. 2011. "Applying for and Staying on the Supplemental Nutrition Assistance Program in South Carolina." Contractor and Cooperator Report No. 65, Food Assistance and Nutrition Research Program, U.S. Department of Agriculture, Washington, D.C.

Ribar, David and Mark Wilhelm. 1999. "The Demand for Welfare Generosity." *Review of Economics and Statistics* 81 (1): 96-108.

Ribar, David C. and Carolyn Wolff. 2013. "What Happened to Cash Assistance for Needy Families?" in Robert Rycroft (ed) *The Economics of Inequality, Poverty and Discrimination in the 21st Century*. ABC-CLIO.

Rosenthal, M. B. and R. G. Frank. 2006. "What is the Empirical Basis for Paying for Quality in Health Care?" *Medical Care Research and Review,* 63(2): 135-157.

Rosenthal, M. B., R. G. Frank, Z. Li, and A. M. Epstein. 2005. "Early Experience with Pay-for-Performance: From Concept to Practice." *Journal of the American Medical Association*, 297(14): 1788-1793.

Rowe, J. W. 2006. "Pay-for-Performance and Accountability: Related Themes in Improving Health Care." *Annals of Internal Medicine*, 145(9): 695-699.

Sarma, S., R.A. Devlin, and W. Hogg. 2010. "Physician's Production of Primary Care in Ontario, Canada." *Health Economics*, 19(1): 14-30.

Shafrin, J. 2010. "Operating on Commission: Analyzing how Physician Financial Incentives Affect Surgery Rates." *Health Economics*, 19(5): 562-580.

Sorensen, R. and J. Grytten. 2003. "Service Production and Contract Choice in Primary Physician Servies." *Health Policy*, 66(1): 73-93.

South Carolina Department of Social Services (SC DSS). 2009. *FI Policy Manual*. Columbia, SC: Department of Social Services.

U.S. Census Bureau. 2011a. "Annual Social and Economic Supplement, POV03: People in Families with Related Children Under 18 by Family Structure, Age, and Sex, Iterated by Income-to-Poverty Ratio and Race: 2008." Accessed June 9. http://www.census.gov/hhes/www/cpstables/032009/pov/new03_100_01.htm.

U.S. Census Bureau. 2011b. Current Population Survey, Annual Social and Economic Supplements. Accessed June 8. http://www.census.gov/hhes/www/poverty/data/historical/families.html.

U.S. Census Bureau. 2011c. "Educational Attainment: CPS Historical Time Series." Accessed June 9. http://www.census.gov/hhes/socdemo/education/data/cps/historical/index.html.

U.S. Department of Health and Human Services (U.S. DHHS), Administration for Children & Families. 2003. "Table 1A, TANF Work Participation Rates, Fiscal Year 2002." Last modified September 2. http://www.acf.hhs.gov/programs/ofa/particip/2002/table01a.htm.

U.S. DHHS, Administration for Children & Families. 2004. "Table 1A, TANF Work Participation Rates, Fiscal Year 2003." Last modified October 28. http://www.acf.hhs.gov/programs/ofa/particip/2003/table01a.htm.

U.S. DHHS, Administration for Children & Families. 2006. "Table 1A, TANF Work Participation Rates, Fiscal Year 2004." Last modified August 8. http://www.acf.hhs.gov/programs/ofa/particip/2004/table01a.htm.

U.S. DHHS, Administration for Children & Families. 2007a. "Table 1A, TANF Work Participation Rates, Fiscal Year 2005." Last modified July 31. http://www.acf.hhs.gov/programs/ofa/particip/2005/tab1a.htm.

U.S. DHHS, Administration for Children & Families. 2007b. "Table 1A, TANF Work Participation Rates, Fiscal Year 2006." Last modified August 15. http://www.acf.hhs.gov/programs/ofa/particip/2006/tab1a.htm.

U.S. DHHS, Administration for Children & Families. 2009a. "Table 1A, TANF Work Participation Rates, Fiscal Year 2007." Last modified August 28. http://www.acf.hhs.gov/programs/ofa/particip/2007/tab1a.htm.

U.S. DHHS, Administration for Children & Families. 2009b. "TANF: Total Number of Families, Fiscal and Calendar Year 2008." Last modified November 22. http://www.acf.hhs.gov/programs/ofa/data-reports/caseload/2008/2008_family_tan.htm.

U.S. DHHS, Administration for Children & Families. 2010a. "Characteristics and Financial Circumstances of TANF Recipients, Fiscal Year 2008: Table 5." Last Modified January 17. http://www.acf.hhs.gov/programs/ofa/character/FY2008/tab05.htm.

U.S. DHHS, Administration for Children & Families. 2010b. "Characteristics and Financial Circumstances of TANF Recipients, Fiscal Year 2008: Table 9." Last Modified January 17. http://www.acf.hhs.gov/programs/ofa/character/FY2008/tab09.htm.

U.S. DHHS, Administration for Children & Families. 2010c. "TANF: Total Number of Families, Fiscal and Calendar Year 2009." Last modified May 21. http://www.acf.hhs.gov/programs/ofa/data-reports/caseload/2009/2009_family_tan.htm.

U.S. DHHS, Administration for Children & Families. 2011a. "AFDC Data Archive – Download." Accessed June 9. http://aspe.hhs.gov/ftp/hsp/AFDC/AFDCdownload.shtml.

U.S. DHHS, Administration for Children & Families. 2011b. "Characteristics and Financial Circumstances of TANF Recipients." Accessed June 9. http://www.acf.hhs.gov/programs/ofa/character/.

U.S. DHHS, Administration for Children & Families. 2011c. "Characteristics and Financial Circumstances of TANF Recipients, Fiscal Year 2008: Table 46." Accessed June 9. http://www.acf.hhs.gov/programs/ofa/character/FY2008/tab46.htm.

U.S. DHHS, Administration for Children & Families. 2011d. "Table 1A, TANF Work Participation Rates, Fiscal Year 2008." Last modified May 23. http://www.acf.hhs.gov/programs/ofa/particip/2008/tab1a.htm.

U.S. DHHS, Administration for Children & Families. 2011e. "Table 1A, TANF Work Participation Rates, Fiscal Year 2009." Last modified May 5. http://www.acf.hhs.gov/programs/ofa/particip/2009/tab01a.htm.

U.S. DHHS, Administration for Children & Families. 2011f. "Temporary Assistance for Needy Families (TANF): Data and Documentation." Accessed June 9. http://aspe.hhs.gov/ftp/hsp/tanf-data/index.shtml.

U.S. DHHS, Administration for Children & Families. 2012. *Temporary Assistance for Needy Families (TANF) Ninth Report to Congress*. Washington, DC.

U.S. Department of Labor, Bureau of Labor Statistics. 2011. Current Population Survey, Labor Force Statistics, Series ID LNS14000000. Accessed June 7. http://data.bls.gov/timeseries/LNS14000000.

U.S. General Accounting Office (GAO). 2005. *Welfare Reform: HHS Should Exercise Oversight to Help Ensure TANF Work Participation is Measured Consistently across States*, GAO-05-821. Washington, DC: General Accounting Office.

U.S. General Accounting Office (GAO). 2010. *Temporary Assistance for Needy Families: Implications of Recent Legislative and Economic Changes for State Programs and Work Participation Rates*, GAO-10-525. Washington, DC: General Accounting Office.

Wee, C., R. Phillips, H. Burstin, E. Cook, A. Puopolo, T. Brennan, and J. Haas. 2001. "Influence of Financial Productivity Incentives on the Use of Preventive Care." *The American Journal of Medicine*, 110(3): 181-187.

Yelowitz, Aaron. 1995. "The Medicaid Notch, Labor Supply, and Welfare Participation." *Quarterly Journal of Economics* 110 (4): 909-39.

APPENDIX A

APPENDIX TABLES

Appendix Table A1. Sample Means

	Full Sample	Private Insurance Subsample	Medicare Subsample	Medicaid Subsample
Time spent with patient (minutes)	21.791	21.473	21.08	20.529
	(0.320)	(0.302)	(0.275)	(0.332)
Number of diagnostic/screening services	3.578	3.568	3.648	3.609
	(0.045)	(0.049)	(0.062)	(0.065)
Number of procedures	0.174	0.181	0.194	0.145
	(0.007)	(0.008)	(0.010)	(0.011)
Productivity incentive pay	0.416	0.428	0.424	0.427
	(0.011)	(0.012)	(0.013)	(0.015)
Patient-centered incentive pay	0.226	0.222	0.214	0.225
	(0.010)	(0.010)	(0.011)	(0.013)
Practice profiling incentive pay	0.087	0.085	0.085	0.095
	(0.006)	(0.007)	(0.007)	(0.009)
Average patient age	46.288	41.479	69.472	34.679
	(0.432)	(0.382)	(0.341)	(0.681)
Fraction of male patients	0.420	0.423	0.418	0.355
	(0.005)	(0.007)	(0.008)	(0.011)
Fraction of patients with imputed sex	0.009	0.009	0.008	0.008
	(0.001)	(0.001)	(0.001)	(0.002)
Fraction of black patients	0.116	0.102	0.097	0.204
	(0.004)	(0.005)	(0.006)	(0.010)
Fraction of non-white and non-black patients	0.057	0.050	0.041	0.068
	(0.003)	(0.003)	(0.004)	(0.006)
Fraction of patients with imputed race	0.271	0.268	0.260	0.236
	(0.009)	(0.009)	(0.011)	(0.012)

Fraction of hispanic patients	0.125	0.106	0.086	0.162
	(0.005)	(0.005)	(0.006)	(0.009)
Fraction of patients with imputed ethnicity	0.279	0.286	0.285	0.241
	(0.009)	(0.010)	(0.011)	(0.013)
Fraction of patients expected to pay primarily with public insurance	0.392	-	-	-
	(0.006)			
Fraction of patients expected to pay primarily with means other than private or public insurance	0.054	-	-	-
	(0.004)			
Fraction of patients with primary reason for visit being an acute problem	0.312	0.338	0.252	0.315
	(0.005)	(0.007)	(0.008)	(0.011)
Fraction of patients with primary reason for visit being a chronic problem	0.427	0.398	0.536	0.411
	(0.007)	(0.008)	(0.010)	(0.013)
Fraction of patients with primary reason for visit being a pre-/post-surgery visit	0.084	0.088	0.094	0.066
	(0.004)	(0.005)	(0.006)	(0.007)
Fraction of patients seen before in the medical practice	0.843	0.830	0.876	0.865
	(0.004)	(0.005)	(0.006)	(0.008)
Fraction of patients with imputed values for having been seen before in the practice	0.010	0.008	0.011	0.013
	(0.001)	(0.001)	(0.002)	(0.003)
Fraction of patients for whom the physician is the primary care doctor	0.373	0.354	0.329	0.461
	(0.010)	(0.011)	(0.012)	(0.015)
Physician is an MD	0.935	0.932	0.933	0.922
	(0.006)	(0.006)	(0.007)	(0.008)
Physician is an owner in the practice	0.528	0.562	0.586	0.453
	(0.011)	(0.012)	(0.013)	(0.015)
Physician sees patients on evenings or weekends	0.272	0.267	0.226	0.291
	(0.010)	(0.010)	(0.011)	(0.014)
Physician is accepting new patients	0.963	0.962	0.960	0.972
	(0.004)	(0.005)	(0.005)	(0.005)
Physician specialty is surgical care	0.246	0.263	0.293	0.197
	(0.01)	(0.011)	(0.012)	(0.012)
Physician specialty is medical care	0.276	0.277	0.311	0.231
	(0.01)	(0.011)	(0.012)	(0.013)
Medical practice performs its own lab testing	0.560	0.551	0.531	0.607

	(0.011)	(0.012)	(0.013)	(0.015)
Medical practice has electronic medical records	0.413	0.415	0.406	0.381
	(0.011)	(0.012)	(0.013)	(0.015)
The medical practice is a private practice	0.759	0.798	0.797	0.673
	(0.010)	(0.010)	(0.011)	(0.015)
Medical practice is located in an MSA	0.892	0.888	0.877	0.855
	(0.007)	(0.008)	(0.009)	(0.011)
Medical practice is located in the northeast	0.191	0.191	0.195	0.192
	(0.009)	(0.009)	(0.011)	(0.012)
Medical practice is located in the midwest	0.253	0.253	0.263	0.280
	(0.010)	(0.010)	(0.012)	(0.014)
Medical practice is located in the west	0.220	0.212	0.192	0.197
	(0.009)	(0.010)	(0.011)	(0.012)
Observations	1930	1756	1408	1037

Note: Data are from the 2006-2008 NAMCS. Standard errors are in parentheses.

Appendix Table A2. Discrete Factor Approximation Models by Insurance Type for Time Spent with Each Patient

	Average Time Spent with Each Patient (Minutes)		
	Private Insurance Patients	Medicare Patients	Medicaid Patients
	(1)	(2)	(3)
Productivity incentive pay	-2.071 **	-2.420 **	-4.218 ***
	(0.982)	(1.044)	(0.935)
Patient-centered incentive pay	-0.539	-0.698	-6.101 ***
	(1.398)	(1.638)	(1.023)
Practice profiling incentive pay	-0.234	-0.571	-1.719
	(1.167)	(1.386)	(1.097)
Average patient age	0.032 **	-0.002	0.010
	(0.016)	(0.022)	(0.016)
Fraction of male patients	-1.767 **	-0.635	0.012
	(0.784)	(0.801)	(0.870)
Fraction of patients with imputed sex	-4.280	2.249	-1.881
	(5.307)	(5.842)	(8.733)
Fraction of black patients	-1.155	-2.381 **	-0.809
	(1.208)	(1.185)	(1.008)
Fraction of non-white and non-black patients	2.723	0.626	2.407
	(1.712)	(2.040)	(1.585)
Fraction of patients with imputed race	-0.572	-0.438	-0.297
	(1.017)	(1.094)	(1.141)
Fraction of hispanic patients	-0.909	-1.951	-0.265
	(1.232)	(1.375)	(1.081)
Fraction of patients with imputed ethnicity	-1.205	-1.153	-0.804
	(0.972)	(0.975)	(1.116)
Fraction of patients with primary reason for visit	-1.336	-1.203	-1.910 *

being an acute problem	(1.180)	(1.286)	(1.135)
Fraction of patients with primary reason for visit being a chronic problem	1.884	0.455	-0.509
	(1.194)	(1.187)	(1.149)
Fraction of patients with primary reason for visit being a pre-/post-surgery visit	-1.523	-1.083	-4.586 **
	(1.439)	(1.785)	(2.335)
Fraction of patients seen before in the medical practice	-8.978 ***	-7.247 ***	-7.500 ***
	(0.990)	(1.126)	(1.100)
Fraction of patients with imputed values for having been seen before in the practice	0.045	1.003	3.825
	(3.549)	(4.874)	(4.763)
Fraction of patients for whom the physician is the primary care doctor	0.343	0.679	-0.725
	(0.838)	(0.933)	(0.988)
Physician is an MD	0.054	0.057	1.154
	(0.951)	(1.219)	(1.287)
Physician is an owner in the practice	-1.816 ***	-1.597 ***	-1.633 **
	(0.529)	(0.611)	(0.730)
Physician sees patients on evenings or weekends	0.067	0.456	1.412 *
	(0.612)	(0.701)	(0.725)
Physician is accepting new patients	-0.105	-1.501	0.560
	(1.260)	(1.273)	(2.226)
Physician specialty is surgical care	-0.585	-0.771	-0.590
	(0.954)	(1.076)	(1.264)
Physician specialty is medical care	3.838 ***	3.643 ***	5.630 ***
	(0.926)	(0.961)	(1.107)
Medical practice performs its own lab testing	-0.556	-0.688	-0.743
	(0.511)	(0.592)	(0.760)
Medical practice has electronic medical records	1.171 **	0.876	0.315
	(0.473)	(0.542)	(0.642)
The medical practice is a private practice	0.642	0.405	-0.587

	(0.633)	(0.777)	(0.831)
Medical practice is located in an MSA	-0.528	-0.585	-1.017
	(0.715)	(0.778)	(0.882)
Medical practice is located in the northeast	1.157 *	0.570	1.022
	(0.641)	(0.722)	(0.860)
Medical practice is located in the midwest	-0.302	-1.010	-0.100
	(0.583)	(0.672)	(0.891)
Medical practice is located in the west	1.397 **	1.213	2.232 **
	(0.695)	(0.755)	(0.947)
ρ_0	0.990	1.149	5.795 ***
	(0.966)	(1.103)	(0.653)
Point 1	-1.245	-1.245	-1.245
Point 2	1.096 ***	1.130 ***	0.937 ***
	-0.095	-0.108	-0.12
Weight	0.562 ***	0.666 ***	0.585 ***
	-0.063	-0.076	-0.068
Log Likelihood	-5641.64	-4518.29	-3343.97
Observations	1756	1408	1037

Note: Data are from the 2006-2008 NAMCS. The dependent variable in these models is average time spent with each patient, in minutes. Standard errors are in parentheses. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Appendix Table A3. Discrete Factor Approximation Models by Insurance Type for Number of Diagnostic and Screening Services

	Number of Diagnostic and Screening Services		
	Private Insurance Patients	Medicare Patients	Medicaid Patients
	(1)	(2)	(3)
Productivity incentive pay	-0.179	-0.153	-0.148
	(0.164)	(0.192)	(0.243)
	-0.164	*-0.137*	*-0.135*
Patient-centered incentive pay	-0.137	-0.161	-0.214
	(0.245)	(0.342)	(0.337)
	-0.125	*-0.144*	*-0.194*
Practice profiling incentive pay	-0.018	-0.046	0.208
	(0.206)	(0.263)	(0.269)
	-0.017	*-0.041*	*0.189*
Average patient age	0.035 ***	0.013 ***	0.025 ***
	(0.003)	(0.005)	(0.003)
	0.032	*0.011*	*0.022*
Fraction of male patients	-0.374 ***	-0.160	-0.075
	(0.144)	(0.163)	(0.158)
	-0.342	*-0.143*	*-0.068*
Fraction of patients with imputed sex	-0.464	0.626	-0.849
	(0.814)	(1.260)	(1.660)
	-0.424	*0.559*	*-0.773*
Fraction of black patients	0.486 **	0.609 ***	0.773 ***
	(0.209)	(0.219)	(0.172)
	0.444	*0.544*	*0.703*
Fraction of non-white and non-black patients	0.002	0.665 **	0.702 **
	(0.311)	(0.323)	(0.289)

	0.002	*0.594*	*0.639*
Fraction of patients with imputed race	-0.100	-0.202	-0.159
	(0.168)	(0.228)	(0.228)
	-0.091	*-0.180*	*-0.145*
Fraction of hispanic patients	0.233	0.468 *	0.170
	(0.201)	(0.264)	(0.206)
	0.213	*0.418*	*0.155*
Fraction of patients with imputed ethnicity	-0.237	-0.151	-0.344
	(0.161)	(0.212)	(0.222)
	-0.217	*-0.135*	*-0.313*
Fraction of patients with primary reason for visit being an acute problem	-0.756 ***	-1.213 ***	-0.524 **
	(0.211)	(0.266)	(0.236)
	-0.690	*-1.082*	*-0.476*
Fraction of patients with primary reason for visit being a chronic problem	-1.142 ***	-1.085 ***	-0.752 ***
	(0.200)	(0.232)	(0.240)
	-1.043	*-0.968*	*-0.684*
Fraction of patients with primary reason for visit being a pre-/post-surgery visit	-0.709 **	-1.089 ***	-0.595 *
	(0.286)	(0.334)	(0.320)
	-0.648	*-0.972*	*-0.542*
Fraction of patients seen before in the medical practice	-0.444 **	-0.377	-0.584 ***
	(0.204)	(0.245)	(0.214)
	-0.405	*-0.336*	*-0.531*
Fraction of patients with imputed values for having been seen before in the practice	0.467	-0.032	0.638
	(0.562)	(0.612)	(0.822)
	0.426	*-0.028*	*0.581*
Fraction of patients for whom the physician is the primary care doctor	0.187	0.187	0.144
	(0.156)	(0.207)	(0.201)
	0.171	*0.167*	*0.131*

Physician is an MD	-0.216	-0.152	-0.344
	(0.200)	(0.250)	(0.237)
	-0.197	*-0.136*	*-0.313*
Physician is an owner in the practice	0.097	0.136	-0.055
	(0.097)	(0.121)	(0.142)
	0.088	*0.121*	*-0.050*
Physician sees patients on evenings or weekends	-0.196 *	-0.320 **	-0.099
	(0.107)	(0.142)	(0.146)
	-0.179	*-0.285*	*-0.090*
Physician is accepting new patients	0.164	0.407	-0.085
	(0.222)	(0.282)	(0.331)
	0.150	*0.363*	*-0.077*
Physician specialty is surgical care	-2.115 ***	-2.212 ***	-2.133 ***
	(0.168)	(0.226)	(0.226)
	-1.932	*-1.975*	*-1.940*
Physician specialty is medical care	-0.900 ***	-1.098 ***	-0.708 ***
	(0.160)	(0.207)	(0.217)
	-0.822	*-0.980*	*-0.644*
Medical practice performs its own lab testing	0.674 ***	0.883 ***	0.578 ***
	(0.091)	(0.115)	(0.124)
	0.615	*0.788*	*0.526*
Medical practice has electronic medical records	0.121	0.193 *	0.035
	(0.084)	(0.106)	(0.114)
	0.111	*0.172*	*0.032*
The medical practice is a private practice	0.186	0.375 **	0.184
	(0.116)	(0.151)	(0.153)
	0.170	*0.335*	*0.167*
Medical practice is located in an MSA	0.409 ***	0.443 ***	0.157

	(0.133)	(0.169)	(0.159)
	0.374	*0.395*	*0.142*
Medical practice is located in the northeast	-0.073	0.009	0.063
	(0.119)	(0.155)	(0.175)
	-0.067	*0.008*	*0.057*
Medical practice is located in the midwest	-0.035	-0.023	0.059
	(0.108)	(0.132)	(0.142)
	-0.032	*-0.020*	*0.054*
Medical practice is located in the west	-0.037	-0.141	0.238
	(0.122)	(0.149)	(0.172)
	-0.034	*-0.126*	*0.217*
ρ_0	0.172	0.150	0.098
	(0.164)	(0.223)	(0.234)
Point 1	-1.244	-1.244	-1.244
Point 2	1.105 ***	1.132 ***	1.068 ***
	-0.095	-0.108	-0.128
Weight	0.558 ***	0.653 ***	0.485 ***
	-0.063	-0.076	-0.089
Log Likelihood	-6460.93	-5198.86	-3825.88
Observations	1756	1408	1037

Note: Data are from the 2006-2008 NAMCS. The dependent variable in these models is average number of iagnostic and screening services provided to each patient. Standard errors are in parentheses. Marginal effects on the censored means are in italics * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Appendix Table A4. Discrete Factor Approximation Models by Insurance Type for Number of Procedures

	Number of Procedures Performed		
	Private Insurance Patients	Medicare Patients	Medicaid Patients
	(1)	(2)	(3)
Productivity incentive pay	-0.015	-0.013	0.175
	(0.055)	(0.084)	(0.127)
	-0.004	*-0.003*	*0.040*
Patient-centered incentive pay	-0.044	-0.293 *	0.001
	(0.081)	(0.150)	(0.204)
	-0.013	*-0.072*	*0.000*
Practice profiling incentive pay	0.113 *	0.019	0.236
	(0.066)	(0.113)	(0.187)
	0.033	*0.005*	*0.054*
Average patient age	-0.001	0.004 *	-0.003
	(0.001)	(0.002)	(0.002)
	0.000	*0.001*	*-0.001*
Fraction of male patients	-0.117 **	0.145 **	0.012
	(0.052)	(0.074)	(0.108)
	-0.034	*0.035*	*0.003*
Fraction of patients with imputed sex	0.114	-0.335	0.276
	(0.226)	(0.606)	(0.779)
	0.033	*-0.082*	*0.064*
Fraction of black patients	-0.075	-0.117	0.047
	(0.076)	(0.124)	(0.125)
	-0.022	*-0.029*	*0.011*
Fraction of non-white and non-black patients	-0.341 ***	-0.107	-0.065
	(0.130)	(0.144)	(0.198)

	-0.099	*-0.026*	*-0.015*
Fraction of patients with imputed race	-0.013	0.003	-0.060
	(0.059)	(0.101)	(0.127)
	-0.004	*0.001*	*-0.014*
Fraction of hispanic patients	-0.087	-0.030	-0.061
	(0.073)	(0.130)	(0.140)
	-0.025	*-0.007*	*-0.014*
Fraction of patients with imputed ethnicity	0.058	0.046	0.109
	(0.056)	(0.097)	(0.122)
	0.017	*0.011*	*0.025*
Fraction of patients with primary reason for visit being an acute problem	0.283 ***	0.379 ***	0.152
	(0.078)	(0.120)	(0.146)
	0.082	*0.093*	*0.035*
Fraction of patients with primary reason for visit being a chronic problem	0.067	0.047	0.117
	(0.080)	(0.106)	(0.155)
	0.019	*0.011*	*0.027*
Fraction of patients with primary reason for visit being a pre-/post-surgery visit	0.392 ***	0.522 ***	0.096
	(0.097)	(0.142)	(0.223)
	0.113	*0.128*	*0.022*
Fraction of patients seen before in the medical practice	-0.090	-0.114	-0.057
	(0.064)	(0.106)	(0.147)
	-0.026	*-0.028*	*-0.013*
Fraction of patients with imputed values for having been seen before in the practice	-0.606	0.117	-0.857
	(0.445)	(0.418)	(0.740)
	-0.175	*0.029*	*-0.198*
Fraction of patients for whom the physician is the primary care doctor	-0.014	0.006	-0.068
	(0.053)	(0.091)	(0.123)
	-0.004	*0.001*	*-0.016*

Physician is an MD	-0.104 *	-0.011	-0.116
	(0.059)	(0.100)	(0.140)
	-0.030	*-0.003*	*-0.027*
Physician is an owner in the practice	0.036	0.021	0.177 *
	(0.033)	(0.053)	(0.091)
	0.010	*0.005*	*0.041*
Physician sees patients on evenings or weekends	0.034	0.090	-0.015
	(0.036)	(0.062)	(0.087)
	0.010	*0.022*	*-0.004*
Physician is accepting new patients	0.018	-0.112	0.167
	(0.070)	(0.116)	(0.214)
	0.005	*-0.027*	*0.038*
Physician specialty is surgical care	0.270 ***	0.411 ***	0.218
	(0.058)	(0.100)	(0.147)
	0.078	*0.100*	*0.050*
Physician specialty is medical care	0.066	0.216 **	-0.314 **
	(0.055)	(0.092)	(0.145)
	0.019	*0.053*	*-0.073*
Medical practice performs its own lab testing	0.039	0.027	-0.033
	(0.032)	(0.055)	(0.088)
	0.011	*0.006*	*-0.008*
Medical practice has electronic medical records	0.046	-0.024	-0.012
	(0.029)	(0.046)	(0.072)
	0.013	*-0.006*	*-0.003*
The medical practice is a private practice	0.081 **	0.073	-0.202 *
	(0.040)	(0.072)	(0.109)
	0.023	*0.018*	*-0.047*
Medical practice is located in an MSA	-0.059	-0.112	-0.076

	(0.042)	(0.071)	(0.104)
	-0.017	*-0.027*	*-0.017*
Medical practice is located in the northeast	-0.015	-0.172 **	0.157
	(0.040)	(0.068)	(0.119)
	-0.004	*-0.042*	*0.036*
Medical practice is located in the midwest	0.035	-0.016	0.173 *
	(0.036)	(0.060)	(0.097)
	0.010	*-0.004*	*0.040*
Medical practice is located in the west	0.121 ***	0.020	0.334 ***
	(0.040)	(0.066)	(0.108)
	0.035	*0.005*	*0.077*
ρ_0	0.017	0.124	-0.181
	(0.054)	(0.092)	(0.137)
Point 1	-1.242	-1.242	-1.242
Point 2	1.106 ***	1.147 ***	1.080 ***
	-0.096	-0.11	-0.128
Weight	0.555 ***	0.691 ***	0.493 ***
	-0.063	-0.074	-0.086
Log Likelihood	-5257.96	-3876.61	-2596.82
Observations	1756	1408	1037

Note: Data are from the 2006-2008 NAMCS. The dependent variable in these models is the average number of procedures performed on each patient. Standard errors are in parentheses. Marginal effects on the censored means are in italics * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Appendix Table A5. Unrestricted Discrete Factor Approximation Model

	Average Time Spent with Each Patient (Minutes)
Productivity incentive pay	-1.728 ***
	(0.637)
Patient-centered incentive pay	-2.534
	(1.921)
Practice profiling incentive pay	-0.136
	(1.044)
Average patient age	0.024 *
	(0.014)
Fraction of male patients	-1.175
	(0.834)
Fraction of patients with imputed sex	0.860
	(4.991)
Fraction of black patients	-0.183
	(1.078)
Fraction of non-white and non-black patients	4.134 ***
	(1.316)
Fraction of patients with imputed race	0.144
	(0.913)
Fraction of hispanic patients	-1.884
	(1.295)
Fraction of patients with imputed ethnicity	-1.830 **
	(0.849)
Fraction of patients expected to pay primarily with public insurance	-1.610 **
	(0.803)
Fraction of patients expected to pay primarily with means other than private or public insurance	-1.133
	(1.506)
Fraction of patients with primary reason for visit being an acute problem	-1.646
	(1.079)
Fraction of patients with primary reason for visit being a chronic problem	2.161 **
	(1.048)
Fraction of patients with primary reason for visit being a pre-/post-surgery visit	-2.281
	(1.569)
Fraction of patients seen before in the medical practice	-9.528 ***
	(1.105)
Fraction of patients with imputed values for having been seen before in the practice	2.008
	(3.660)
Fraction of patients for whom the physician is the	0.891

primary care doctor	(0.770)
Physician is an MD	0.245
	(0.985)
Physician is an owner in the practice	-2.222 ***
	(0.513)
Physician sees patients on evenings or weekends	0.382
	(0.550)
Physician is accepting new patients	-1.119
	(1.186)
Physician specialty is surgical care	0.008
	(0.943)
Physician specialty is medical care	4.312 ***
	(0.849)
Medical practice performs its own lab testing	-0.782
	(0.504)
Medical practice has electronic medical records	0.912 **
	(0.453)
The medical practice is a private practice	-0.016
	(0.591)
Medical practice is located in an MSA	-0.985
	(0.725)
Medical practice is located in the northeast	1.002
	(0.630)
Medical practice is located in the midwest	-0.525
	(0.581)
Medical practice is located in the west	1.539 **
	(0.643)
ρ_0	2.253 *
	(1.361)
ρ_1	1
ρ_2	2.190 ***
	(0.319)
ρ_3	1.323 ***
	(0.190)
Point 1	-1.245
Point 2	0.489 ***
	(0.113)
Weight	0.671 ***
	(0.050)
Log Likelihood	-6190.30

Note: Data are from the 2006-2008 NAMCS. There are 1,930 observations. Standard errors are in parentheses. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Appendix Table A6. Discrete Factor Approximation Models by Health Status for Time Spent with Each Patient

	Time Spent with Patient	
	No Chronic Conditions	One or More Chronic Conditions
	(1)	(2)
Productivity incentive pay	-3.559 ***	-2.147 **
	(0.890)	(0.993)
Patient-centered incentive pay	-2.249 **	-1.248
	(1.011)	(1.312)
Practice profiling incentive pay	-1.317	-0.416
	(1.022)	(1.169)
Average patient age	0.015	0.001
	(0.014)	(0.014)
Fraction of male patients	-0.591	-0.532
	(0.721)	(0.752)
Fraction of patients with imputed sex	-8.299	2.379
	(9.760)	(4.270)
Fraction of black patients	-1.690	-1.072
	(1.104)	(1.025)
Fraction of non-white and non-black patients	2.307	1.988
	(1.522)	(1.254)
Fraction of patients with imputed race	-0.285	0.522
	(0.962)	(0.899)
Fraction of hispanic patients	-1.776	-2.301 **
	(1.168)	(1.098)
Fraction of patients with imputed ethnicity	-0.794	-2.054 **
	(0.968)	(0.838)

Fraction of patients expected to pay primarily with public insurance	-0.494	-1.706 **
	(0.783)	(0.756)
Fraction of patients expected to pay primarily with means other than private or public insurance	-1.156	-1.533
	(1.354)	(1.431)
Fraction of patients with primary reason for visit being an acute problem	-1.778 *	-3.159 ***
	(0.998)	(1.024)
Fraction of patients with primary reason for visit being a chronic problem	-0.809	0.588
	(1.078)	(0.992)
Fraction of patients with primary reason for visit being a pre-/post-surgery visit	-2.998 **	-1.834
	(1.497)	(1.614)
Fraction of patients seen before in the medical practice	-7.898 ***	-8.955 ***
	(0.898)	(0.998)
Fraction of patients with imputed values for having been seen before in the practice	6.643	-0.782
	(4.360)	(4.142)
Fraction of patients for whom the physician is the primary care doctor	-0.883	1.941 **
	(0.829)	(0.768)
Physician is an MD	-0.016	0.318
	(0.955)	(1.053)
Physician is an owner in the practice	-2.666 ***	-1.657 ***
	(0.512)	(0.523)
Physician sees patients on evenings or weekends	0.399	0.254
	(0.597)	(0.542)
Physician is accepting new patients	-1.207	-0.640
	(1.037)	(1.254)
Physician specialty is surgical care	-0.128	-0.106
	(0.921)	(0.980)
Physician specialty is medical care	4.655 ***	4.929 ***
	(0.888)	(0.852)

Medical practice performs its own lab testing	-0.198	-0.967 *
	(0.519)	(0.511)
Medical practice has electronic medical records	1.243 ***	1.074 **
	(0.467)	(0.464)
The medical practice is a private practice	0.171	-0.286
	(0.616)	(0.617)
Medical practice is located in an MSA	-0.674	-0.680
	(0.744)	(0.711)
Medical practice is located in the northeast	0.213	1.088 *
	(0.631)	(0.643)
Medical practice is located in the midwest	-0.882	-0.676
	(0.615)	(0.593)
Medical practice is located in the west	0.961	1.331 **
	(0.672)	(0.640)
ρ0	2.686 ***	1.262
	(0.697)	(0.920)
Point 1	-1.245	-1.245
Point 2	1.114 ***	1.138 ***
	-0.097	-0.094
Weight	0.504 ***	0.581 ***
	-0.058	-0.062
Log Likelihood	-5477.27	-5910.12
Observations	1718	1847

Note: Data are from the 2006-2008 NAMCS. Standard errors are in parentheses. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.