Essays on Health Economics

Wilfredo A. Lim, Jr.

Submitted in partial fulfillment of the
requirements for the degree
of Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2012

ABSTRACT

Essays on Health Economics

Wilfredo A. Lim, Jr.

This dissertation consists of three essays on health economics. The first chapter provides empirical evidence on the impacts of government reimbursement of long-term care. We apply a regression discontinuity design using administrative data from South Korea to estimate the impact of subsidies for formal home and institutional care on informal care use and medical expenditures. We find that reimbursement leads to increases in formal long-term care utilization, even accounting for crowd out of private spending. Among individuals who are partially dependent for some activities of daily living (ADLs), we find that increased use of formal home care has no impact on the use of informal care at the extensive margin or on medical expenses. Among individuals who are partially dependent for several ADLs, we find that increased use of institutional care leads to reductions in informal care and medical expenses. Among individuals who are completely dependent for several ADLs, we find that substitution of home care for institutional care leads to substantial decreases in medical spending.

The second chapter studies state laws passed in the late 1990s that required health insurers to cover diabetes related equipment, supplies, and education. We assess the impact of these mandates on health related behavior and labor market outcomes. We find no significant effects for diabetics or groups with higher prevalence of diabetes in terms of exercise, diet, income, or employment. These results are robust to different specifications and datasets.

The third chapter provides empirical evidence on both outcomes and potential mechanisms resulting from information obtained from screening. We apply a regression discontinuity design using administrative data from South Korea to estimate the impact of different classifications of overall health that vary discontinuously with blood sugar level. We find that secondary examinations due to a "disease suspected" classification leads to follow-up rates greater than 50%. However, we find few impacts otherwise, including short and medium run medical activity and longer run health outcomes. We also find that the responsiveness to the classifications among the highest income quintiles is lower than among the other quintiles, consistent with more educated individuals incorporating information directly from the blood sugar measure itself.

Table of Contents

List of Figures

List of Tables

Acknowledgments

This dissertation has benefitted from discussions with and comments from Janet Currie, Olle Folke, Tal Gross, Wojciech Kopczuk, Benjamin Marx, Christine Pal, Miguel Urquiola, Till von Wachter, and seminar participants at Columbia University. For support of my research, I thank Columbia University, the Sasakawa Young Leaders Fellowship Fund, Yonsei University, and the National Health Insurance Corporation of Korea.

I especially thank my advisors who have provided much guidance and support during my dissertation. Doug Almond, Bentley MacLeod, and Kiki Pop-Eleches have provided much feedback, support, time, and patience, all of which have made me into a better economist. Serena Ng has always supported and guided me throughout this program.

This dissertation would not have been possible without my friend and colleague, Hyuncheol Kim. I am indebted to your inspiration, encouragement, and selflessness.

I thank my classmates and friends who have supported me through the ups and downs of graduate school and living in New York City. You have made the five years so enjoyable. I am grateful to my family, who has always been supportive and loving in whatever I have pursued. To Andrea, thank you so much for your love, support, patience, and encouragement. I could not have done this without you.

Most of all, I thank and praise God, the true invisible hand from whom all good things come.

to Andrea

my love and best friend

and to our parents

who have given everything for us

Chapter 1

Formal Long-Term Care Subsidies, Informal Care, and Medical Expenditures

with Hyuncheol Kim

1.1 Introduction

As countries face rapidly aging populations and rising healthcare costs, policies affecting long-term care—services targeting health or personal needs for people with chronic illness or disability—become increasingly important. For example, the share of those age 65 and over in the United States is expected to increase from 13.0% in 2010 to 20.2% in 2050. For Korea, the corresponding shares are 16.5% and 38.2%.[1] Moreover, the shares of those age 80 and over, for whom the need for long-term care is highest, are expected to double from 3.7% to 7.4% in the United States and increase severalfold from 1.9% to 14.5% in Korea.[2] At the same time, societal changes such as declining family size and rising female labor force participation are likely to reduce the availability of family caregivers. In terms of costs, public and private spending on long-term care in the U.S. totaled $183 billion in 2003, or 1.6% of GDP (GAO (2005)). Moreover, a third of Medicaid spending in 2006 went towards long-term care (CBO (2007)).

Much of long-term care is provided informally. As needs expand and costs rise, understanding the role of informal care in meeting this escalating demand becomes increasingly important. This paper aims to shed light on an important aspect—the substitutability of formal for informal care. For example, if formal long-term care services directly substitute for—rather than supplement—informal care, the cost of provision will rise without necessarily increasing the total care received by disabled persons. This could have welfare consequences for the caregivers in terms of their participation in the labor force as well as on intergenerational household bargaining. Thus, understanding the welfare impacts will require understanding under what situations and through which services formal care substitutes for informal care. Additionally, as governments develop and refine long-term care

---

[1]Data are from http://dx.doi.org/10.1787/888932400893, in Colombo et al. (2011).

[2]Data are from http://dx.doi.org/10.1787/888932400874, in Colombo et al. (2011).

policies, implications for economic efficiency will be substantial. Informed policies will need to assess the costs and benefits of subsidizing various types of care—in particular, home versus facility—measured both by direct costs of subsidization as well as potential costs or savings from other medical expenses.

In this paper, we study subsidies for formal home and facility care and their corresponding impacts on informal caregiving and medical expenditures in Korea. This study has a number of advantages that allow us to address this topic and improve upon the existing literature. First, we account for endogeneity in the choice of long-term care by using plausibly exogenous variation induced by a regression discontinuity design. Specifically, long-term care benefits in Korea are assigned based on an assessment score that is very difficult to precisely control. Second, these benefits vary at multiple cutoffs which allow us to separate the impact of home and institutional care benefits. Specifically, the first set of thresholds isolates the impact of just home care benefits; the second set of thresholds allows us to compare home or institutional care benefits versus just home care benefits; and the third set of thresholds allows us to look at the impact of an increase in the relative price of institutional care. Third, our analysis benefits from unique administrative data on formal home and institutional care, informal care, and medical expenditures.

Our main finding is that the benefits of home and facility care are heterogeneous across physical function level and therefore setting policy appropriately has the potential to dramatically reduce medical expenses. Specifically, substantial reductions in medical expenses arise from incentivizing transitions from home to facility care among people who are partially dependent for several activities of daily living, while incentivizing transitions from facility to home care among people who are completely dependent for several ADLs. This finding is not likely culturally or context specific and is consistent with programs in the U.S. such as Money Follows the Person that seeks to transition people with Medicaid from institutions to the community. We also find that formal long-term care is not a strong substitute for informal long-term care at the extensive margin, but find evidence that it does so at the intensive margin. Indeed, given that Korea is a "strong family ties" country, we argue that our results constitute a lower bound for similar effects in the U.S.

Specifically, we find that among individuals who are partially dependent for some activities of daily living[3], government subsidies for formal home care lead to an increase in its utilization, with no impact on informal caregiving at the extensive margin, as measured by child caregiving and independent living. We do find evidence for a reduction at the intensive margin, measured by the use of short-term respite care, which provides temporary relief for the receipient's caregiver. We also find no impact on medical expenses. Among individuals who are partially dependent for several activities of daily living, reimbursement of institutional care leads to an increase in its utilization with corresponding reductions in informal caregiving and medical expenses. Among individuals who are completely dependent for several activities of daily living, we find that an increase in the relative price of institutional care leads to substitution of home care for institutional care. We find no impact on informal caregiving, but we find substantial decreases in medical spending, largely accounted for by a reduction in hospital expenses. From a policy perspective, these findings suggest that among more able individuals, home care may be reduced with minimal detriment to their health; and that among the less able, incentives to transition from facility to home care may improve quality of life and reduce program costs overall.

We explore additional mechanisms for explaining our findings. First, we determine whether crowd out explains our lack of findings on informal care. While we find that public long-term care insurance leads to partial crowd out of private spending on long-term care, long-term care utilization still increases overall. Thus, crowd out is not likely the sole reason

[3]Partially dependent for some ADL's roughly corresponds to needing assistance moving around; partially dependent for several ADL's roughly corresponds to being unable to move on one's own; completely dependent for several ADLs roughly corresponds to being bedridden. See Table 1.1.

for our limited findings on informal care. We also assess the impact of long-term care insurance on short run-mortality, as this measure is important in and of itself and in order to rule out differential mortality in affecting our estimates. We find no statistically significant differences in mortality across all thresholds. Lastly, we show that our results are robust to various checks and specifications of our estimation strategy.

The remainder of this paper is structured as follows. Section 1.2 provides a brief discussion of the literature and our contribution. Section 1.3 explains Korea's Long-Term Care Insurance program and motivates our empirical strategy. Section 1.4 describes the data. Section 1.5 provides a conceptual framework for considering the impacts of subsidies for long-term care. Sections 1.6 and 1.7 present the empirical framework and results, followed by additional robustness checks in Section 1.8. Section 1.9 provides a brief discussion and Section 1.10 concludes.

1.2 Literature Review

Much of the empirical work on understanding the substitutability of formal for informal care is limited in scope and suffers from endogeneity concerns (papers that do not account for endogeneity include Soldo (1985), Wolinsky, Mosely, and Coe (1986), and Bass and Noelker (1987)). The concern with endogeneity is that absent an exogenous source of variation, confounding unobserved characteristics as well as the joint nature of the formal versus informal care decision may lead to misleading findings. For example, to the extent that formal and informal care are positively correlated with unobserved negative health shocks, a naive analysis would find them to be complements even if they were substitutes. Also problematic is that they may be substitutes in some instances and complements in others. For example, an individual may rely on a child caregiver for assistance with basic activities of daily living but may seek formal assistance for more skill-intensive needs such as physical therapy. This highlights the importance of being able to address endogeneity as well as account for different types of formal care and informal care.

One way to address the issue of endogeneity is through the use of instrumental variables. Using number of adult children and presence of a daughter who has no child at home as instruments, Lo Sasso and Johnson (2002) find that frequent help from children for basic personal care reduces the likelihood of future nursing home use. Using number of children and whether the eldest child is a daughter as instruments, Van Houtven and Norton (2004) find that informal care reduces home health care and nursing home use. Using children's gender, marital status, and distance as instruments, Charles and Sevak (2005) find that receipt of informal home care reduces the probability of future nursing home use. However, it is unclear whether the necessary exclusion restrictions would be satisfied, given the complexity of fertility decisions and bargaining over intergenerational transfers. Thus, it is useful to assess the robustness of these results through studies based on more plausibly exogenous sources of variation.

The Balanced Budget Act of 1997 induced such a source of variation. This act led to regional variation in overall decreases in Medicare reimbursement for home care services. Using this source of variation, McKnight (2006) finds resulting reductions in home care utilization that were not offset by increases in institutional care or other medical care. She also finds no adverse health consequences as a result of the policy. Using the same source of variation, Orsini (2010) and Engelhardt and Greenhalgh-Stanley (2010) find reductions in independent living, and Golberstein et al. (2009) find increases in the probability of the use of informal caregiving. A significant limitation these papers share, however, is that due to data limitations and their source of identification, they focus primarily on the provision of home care.

The Channeling demonstration in the U.S. provides another opportunity to assess the relationship between informal and formal care, through randomized evaluation. This experiment sought to substitute a system of home and community care for institutional care. Christianson (1988) and Pezzin, Kemper, and Reschovsky (1996) assess the impact of public home care provision and find limited reductions in the care provided by informal caregivers. However, the latter paper does find a significant increase in the probability that unmarried persons live independently. This highlights the importance of considering both informal caregiving directly and independent living. The results of the Channeling demonstration are limited, however, in that the sample population was particularly frail and the scope was inherently limited to the provision of home care, not institutional care.

Regarding impacts on other medical expenditures, McKnight (2006) finds that reductions in home health care reimbursement and utilization did not lead to increases in other medical care and were not associated with adverse health consequences. Evaluating the impact of Channeling on other medical expenses, Wooldridge and Schore (1988) find large reductions in nursing home use among those who were already in a nursing home at baseline but no impact on the use of hospital, physician, and non-physician medical services.

Our view of the literature is that evidence on the substitutability of institutional and informal care is very limited and is based mostly on observational studies. Moreover, even though understanding the impact of institutional care on health and other medical expenses is necessary for cost-benefit analyses, very little is known at this point.[4] In addition, while there is more work on the substitutability of home and informal care, this evidence is limited in accounting for institutional care and in being generalizable to a broader population of the elderly. This study attempts to fill these gaps directly. By using longitudinal administrative data with measures of home care, institutional care, informal care, and medical expenditures, and a unique policy affecting a broader population of the elderly, we are able to account

[4] In a review paper, Ward et al. (2008) conclude "there is insufficient evidence to compare the effects of care home environments versus hospital environments or own home environments on older persons rehabilitation outcomes."

for the complex interrelationship among informal, home, and institutional care, as well as evaluate the corresponding impacts on health and medical expenses.

Lastly, much of the literature is based on findings in the United States and other Western countries. Other studies outside the U.S. include Stabile et al. (2006) for Canada and Bolin et al. (2008) for Europe. This paper contributes to this literature by providing evidence from an Asian country.

1.3 Background

Korea implemented universal health coverage in 1989. Individuals are covered either by National Health Insurance (NHI) or Medical Care Assistance (MCA), though both programs are overseen by the National Health Insurance Corporation (NHIC). The primary distinction between NHI and MCA is that the latter serves poor individuals. While health insurance coverage includes outpatient care, inpatient care, and prescription pharmaceuticals, no coverage for long-term care is included. In response to this, and due to the demographic and cultural changes affecting the need and provision of long-term care, National Long-Term Care Insurance was implemented in July 2008. This provides coverage for individuals age 65 and over and those with age-related needs such as dementia and Parkinson's disease.

1.3.1 Benefits

Long-term care insurance covers two categories of service benefits: home care and institutional care.[5] Home care includes services provided at the beneficiary's residence. This includes home help where a caregiver provides support for physical activities or housework, home bathing where a caregiver assists the beneficiary in bathing, and home nursing where

[5]In exceptional cases (e.g. for individuals who live in remote regions with no access to long-term care services), cash benefits are provided. However, this represents less than 0.2% of cases.

a nurse provides assistance with such things as medication and dental hygiene. Also included within home care benefits is short-term respite care which covers a short-term stay in a facility to allow the caregiver relief from caregiving activities. Lastly, equipment for the support of daily tasks and physical activities (e.g. a wheelchair) is also included in home care benefits. Institutional care benefits cover long-term residence in a facility where meals, care, and other necessities required for daily function are provided. See Table B.1 for more details. As in the case for general health care, the delivery of long-term care is primarily administered through private providers.

1.3.2 Eligibility

To receive long-term care benefits, individuals must apply, submit a doctor's referral, and be evaluated by an assessment team from the NHIC. Benefits are determined based on an adjusted score, which is the sum of two components, a preliminary score and committee points. The preliminary score is a complex, highly nonlinear function of the responses to 52 evaluation questions, encompassing physical and cognitive function, behavior, nursing assistance, and rehabilitation.[6] Then a local assessment committee, following guidelines determined at the national level, is able to add or subtract up to five points to this score, based on the assessment questions and the doctor's referral.[7]

The adjusted score is used to determine benefits, as depicted in Table 1.1. Individuals who score below 55 are not eligible for long-term care benefits. Individuals who score 55 or above (Grade 3) are eligible for reimbursement of formal home care services up to 740 USD per month, which corresponds to approximately two hours of home help care per

[6] An example of a physical function question is whether the individual is fully independent, partially dependent, or fully dependent for bathing. For more details, including calculation of the preliminary score, see Appendix A.

[7] Committee members are trained annually and when the guidelines are changed.

day.[8] Individuals who score 75 or above (Grade 2) become eligible for reimbursement of institutional care or a home care benefit maximum of 880 USD per month.[9] Individuals who score 95 or above (Grade 1) continue to be eligible for reimbursement of institutional care or a home care benefit maximum of 1040 USD per month. The price of institutional care is 40 USD per day and 45 USD per day for individuals in Grades 2 and 1, respectively. To the extent that there is a copayment, this implies that the cost of institutional care for an individual scoring 95 is discretely higher than the cost for an individual scoring 94.9. As a result, the increased cost of facility care along with the more generous home care benefit incentivizes individuals to transition from institutional to home care at the margin.

Applicants are notified of their classification, not their score. They are reevaluated when major changes to their physical or mental status occur, for the renewal of benefits, or if they appeal for a reevaluation.[10] Benefits must be renewed every twelve months, with the exception of those with significantly high scores (> 100) who may have up to eighteen months.

Figure 1.1 illustrates the committee component of the score in relation to the preliminary score. Note first that most activity occurs within 5 points of the actual thresholds (55, 75, 95).[11] Focusing on preliminary scores in the range [50,55) we see that some individuals are given enough points so that their adjusted scores exceed 55, leading to eligibility for Grade 3 home only benefits. It appears that points are rarely added or subtracted unless doing so changes the eligibility status. Focusing on scores just above 55, the number of instances

[8] See Table 1.1 for general descriptions of individuals falling into each category. All amounts in this paper are converted to USD at the rate of 1100 KRW : 1 USD.

[9] If one were to use both types of care in the same month, the home care benefit would be prorated based on the number of facility days used. However, home and facility care are inherently incompatible with each other (in our data, only 3% of individuals utilize both types of benefits in the same year). Thus, the use of both types of services in the same month is likely due to changes in health status as opposed to joint use.

[10] They are able to appeal indefinitely, though this process typically takes longer than one month.

[11] In practice, scores outside of five points from a threshold are less likely to be reviewed by the committee.

where points are deducted is negligible. Focusing on scores below 50, we see that the number of instances where points are added is negligible, reflecting the fact that any additional points less than 5 would not be enough to become eligible. We find similar patterns in committee action around the remaining thresholds, except we see more instances of subtracted points.

Figure 1.2a illustrates from another perspective how the committee component of the score influences eligibility around the 55 threshold. It also highlights the source of identification in our research design. The probability that the adjusted (post-committee) score exceeds the 55 point threshold is plotted against the preliminary (pre-committee) score.[12] When the preliminary score is below 50, the probability that the adjusted score exceeds the 55 point threshold is effectively zero, consistent with the guideline that the maximum number of points that can be added is five. When the preliminary score is above 55, the probability that the adjusted score exceeds the 55 point threshold is effectively one, reflecting the rarity with which the committees subtract points around this threshold. Between 50 and 55, enough points are added to the preliminary scores of a fraction of individuals so that their adjusted scores exceed 55. Note that this illustration suggests an implicit threshold at 50 (and similarly at 70 and 90). That is, scores above the explicit threshold of 55 virtually guarantee eligibility; scores below the implicit threshold of 50 virtually exclude the possibility of eligibility.

Correspondingly, this figure illustrates the source of identification for our analysis—namely, comparing similar individuals who have different probabilities of treatment.[13] For instance, those with preliminary scores just below 50 have a probability of eligibility for home care benefits of zero. Those with preliminary scores just above 50 have a probability of about 8 percent. This allows us to use variation in the probability of eligibility in order

[12]See Section 1.6 for a discussion of the specification used to generate the figures.

[13]We discuss our empirical strategy more formally in Section 1.6.

to look at the impact of eligibility on reimbursed formal long-term care utilization and relevant outcomes, including independent living, informal caregiving, and medical expenditures. Moreover, the different grades of benefits afford us the possibility of studying several aspects of long-term care utilization. The 50 and 55 thresholds isolate the impact of home care benefits, while the 70 and 75 thresholds isolate the impact of home and institutional care benefits versus just home care benefits. The 90 and 95 thresholds allow us to look at the impact of an increase in the price of institutional care along with an increase in the maximum benefit for home care.

1.3.3 Financing

Long-term care insurance is financed by the government (20%), copayments (up to 20%), and insurance contributions. Insurance contributions were 0.21%, 0.24%, and 0.35% of wages in 2008, 2009, and 2010, respectively. Employers paid 50% of this amount. The copayment for home care services is 15% while that of institutional care is 20%, but the poor (MCA individuals) are exempt from copayments, and individuals with certain conditions faced reduced copayments.[14]

1.4 Data

This study uses a merged dataset combining NHIC administrative data for National Long-Term Care Insurance (NLTCI) and National Health Insurance (NHI). The sample consists of 171,373 individuals who were assessed in 2008 and 2009. The NLTCI data spans 2009 and the first half of 2010 and contains information on gender, age, living and caregiving arrangements, preliminary and adjusted scores from the first eligibility assessment, and reimbursed long-

[14]Individuals who face reduced copayments include the disabled, people with rare and incurable diseases, and the marginally poor.

term care utilization.[15] The NHI data spans 2007 through 2009 and contains annual total, hospital, outpatient, and pharmacy expenditures. Our main explanatory variable is the 2009 preliminary score. Our main measures of formal care are reimbursed home care expenditures and number of institutional care days. We measure home care in expenditures as an aggregate measure to capture the variety of home care services that are used. Our main measures of informal care are an indicator of whether a child is the primary caregiver and an indicator of whether the individual lives alone or with a spouse. The latter measure is our measure of independent living, consistent with the previous literature. Our main measures of medical expenditures are total medical and hospital expenses.[16]

Table 1.2 displays summary statistics by grade. All measures are at baseline, except for long-term care facility days and home care expenditures. ADL Index is a composite score based on activities of daily living questions from the assessment, with a higher number indicating less function. Individuals with lower grades are sicker as measured by the ADL Index, medical expenditures, and hospital days, and tend to have more resources as measured by insurance contribution and MCA percentage. Finally, sicker individuals are less likely to have a child caregiver and live independently.

[15]Because we only observe NLTCI data for the first half of 2010, our sample is reduced by approximately half when looking at informal care outcomes. Analysis of predetermined variables shows covariates are balanced in the reduced sample.

[16]These amounts are inherently exclusive of long-term care expenses.

1.5 Conceptual Framework

1.5.1 Household Responses to Public Long-Term Care Reimbursement

We adapt the model developed by Stabile, Laporte, and Coyte (2006) in order to determine what implications arise from public reimbursement for long-term care.

Consider a two person household consisting of an elderly care recipient and an informal caregiver (e.g. a child). Let household utility be

$$U(X, L, A)$$

where X represents market goods and services, L the leisure time of the caregiver, and A the care recipient's functional ability. The care recipient's ability is defined by the technology

$$A = A(C, H, F)$$

where C is time spent delivering informal care, H is formal home care, and F is institutional (facility) care. Time and financial constraints are satisfied if

$$P_X X + P_H(1 - s_H)H + P_F(1 - s_F)F + WC = V + W(T - L)$$

where P_X is the cost of market goods and services, P_H is the cost of formal home care, P_F is the cost of facility care, s is the relevant government subsidy (in other words, 1-copay), V is non-wage income, W is the cost of the caregiver's time, and T is the total time for leisure, caregiving, and labor market work. The household selects performance ability A so that the marginal benefit of greater ability is equal to the marginal cost of its production. The household cost-effectively selects H, F, and C in order to achieve ability A. L is selected so that the marginal benefit of leisure equals the marginal cost of foregone market goods and services.

We now illustrate the relevant intuition and predictions derived from the model (see Stabile, Laporte, and Coyte (2006) for a more extensive treatment). When an individual is ineligible for reimbursed benefits, she may pay out-of-pocket for H at price P_H. Grade 3 benefits provide a subsidy for H, reducing its effective price to $P_H(1 - s_H)$ up to the maximum level of benefits m_H. This is depicted in Figure 1.4, where the isocost line rotates out as the price of H falls from P_H to $P_H(1 - s_H)$, up to the point where $H = m_H$. After this point, the price returns to P_H. Through an income effect, these benefits will increase the optimal level of A and lead to corresponding increases in C and H if these are normal inputs to its production. Because H is cheaper relative to C, the substitution effect will lead to increases in H but decreases in C. Thus, while Grade 3 benefits are predicted to lead to increases in A and H, the net impact on C is unclear.

Grade 2 benefits lead to both an increase in the maximum level of home benefits, m_H, as well as provide a subsidy for facility benefits, s_F. Note that home and facility care are effectively perfect substitutes in the production of A as they are inherently incompatible with each other.[17] This is reflected in Figure 1.5, where the isocost line rotates out as the effective price of F falls, and the individual chooses to utilize only F. To the extent that F and C are substitutes, this should lead to an increase in F and decreases in C and H. If the individual decides not to utilize F, then the impact of m_H on H would depend on the amount used with only Grade 3 benefits, as in shown in Figure 1.6. If the individual were using less than the maximum beforehand, there would be no impact on A, C, or H. If the individual were using the maximum, this would lead to a pure price effect, resulting in an

[17] In our data, only 3% of individuals utilize both home and facility benefits in the same year, and this is likely due to changes in health status as opposed to joint use.

increase in A and H, but a decrease in C. Therefore, we expect A and F to increase, but C to decrease. The impact on H is uncertain.

Grade 1 benefits lead to both an increase in the maximum level of home benefits, m_H, as well as an effective increase in the price of facility benefits, P_H, as discussed in Section 1.3.2. Thus, the impact of these benefits is a combination of the figures for previous benefits. We expect the increase in the relative price of F to entice some people to switch from F to H (reverse of Figure 1.5). Combined with an increase in m_H (Figure 1.6) we expect a decrease in F and increase in H. The impact on A is ambiguous, however, as the impact of the relative increase in P_F may not be offset by the increase in m_H. The impact on C is also ambiguous and depends again on whether H and C are substitutes or complements.

In summary, the model yields the following predictions for government reimbursement of long-term care:

1. Grade 3 benefits lead to an effective price decrease in home care. As a result, we expect increases in ability and home care. The impact on informal caregiving will depend on whether home care and informal caregiving are substitutes or complements.

2. Grade 2 benefits lead to an effective price decrease in facility care and an increase in the maximum level of home care benefits. Thus, we expect increases in ability and facility care, and a decrease in informal caregiving. The impact on home care is ambiguous.

3. Grade 1 benefits lead to an effective price increase in facility care and an increase in the maximum level of home care benefits. Thus, we expect an increase in home care and a decrease in facility care. The impacts on ability and informal care is ambiguous.

1.6 Empirical Framework

We conduct a regression discontinuity analysis at the thresholds 50, 55, 70, 75, 90, and 95 of the preliminary score that exploit the discontinuous probabilities of eligibility resulting from the committee adjustment portion of the score. Specifically, the aim is to compare outcomes across individuals with similar characteristics but differing probabilities of eligibility for benefits.

The corresponding regression model we estimate is:

$$\text{outcome} = \beta \mathbb{1}\{S \geq \tau\} + f(S) + \gamma X + \epsilon, \tag{1.1}$$

where S is the preliminary score, $f(S)$ is a function of the score, τ is the relevant cutoff, and X is a set of control variables—age, gender, insurance dummies, region type dummies, and health insurance contribution (a proxy for income)—which serve to improve precision of the estimates.

In implementing the regression discontinuity design, an important consideration is the modeling of $f(S)$. One approach is to model it parametrically through linear, quadratic, or higher order polynomials that are allowed to differ on each side of the cutoff. The other approach, which we follow here, is to estimate the discontinuity nonparametrically, which we implement by local linear regression with a rectangular kernel.[18] Our preferred estimates are based on a bandwidth of 2.5 points, in order to reduce bias by staying close to the cutoff while still maintaining enough precision. To assess the sensitivity of our results, we also present results from the optimal bandwidth determined by the procedure in Imbens and Kalyanaraman (2009), hereafter abbreviated IK. We also evaluate the sensitivity of our

[18] As noted in Lee and Lemieux (2010), the choice of kernel typically has little impact and while a triangular kernel is boundary optimal, a more transparent way of putting more weight on observations close to the cutoff is to reestimate a rectangular kernel based model using a smaller bandwidth.

results to other bandwidths and higher order polynomials in Section 1.8.3.

A critical assumption to our identification strategy is that individuals just below a threshold are indeed comparable to individuals just above a threshold. One potential threat to this assumption is whether individuals are able to precisely sort around the threshold (Lee (2008)). If this assumption holds, then one implication is that the density of scores should be continuous around the threshold. Figure 1.3 displays the density of scores, in 0.1-point bins, in our sample around each threshold. With the exception of 75, we see no indication that the density is discontinuous around the threshold. Figure B.1a displays a smoother density of scores, in 1-point bins, which suggests a discontinuity in the density at 55. To address concerns of possible sorting, Figure B.1b displays the density of scores for those who were assessed in April of 2008, the first opportunity for eligibility evaluations and two months before the actual launch of the program. To the extent that the patterns in the 2009 density are due to sorting, we would not expect to see them in the April 2008 density, when individuals have no experience with how responses are mapped into scores. A comparison of Figures B.1a and B.1b indicates that the distribution of scores around the thresholds is strikingly similar for both periods.

Figure B.2 provides evidence for the complexity of the score function and the amount of variation inherent in the score. We take the set of individuals who responded "fully independent" for changing position and changed their response to "needs partial support." We recalculate their score and then plot this against their original score. Highlighting how highly interactive the score function is, note how the change in the response may lead to a change in the score ranging from a few points to more than ten points. This example indicates three things. First, it is difficult to precisely control the score. Second, there is a large degree of randomness within a few points. Third, it is possible that a response that indicates a sicker individual may actually lead to a *reduction* in points. This results from the highly interactive nature of the way the score is calculated.[19]

To the extent that there is no sorting and that the observed distribution of scores is due to the score function, individuals on each side of the threshold may still be comparable. As discussed in Urquiola and Verhoogen (2009), stacking alone may not violate the regression discontinuity assumptions since violation arises from the interaction of the stacking and the endogenous sorting of individuals. Thus, the more fundamental question for our identification strategy is whether the distribution of predetermined characteristics is identical on each side of the threshold. We show in Section 1.8.1 that with the exception of the 75 threshold, predetermined characteristics appear balanced around each threshold.

1.7 Results

We begin with our main results on the impact of eligibility on reimbursed utilization of formal long-term care, informal caregiving, and medical expenditures in Section 1.7.1. In Section 1.7.2, we address crowd out of private spending on formal-long term care and other potential explanations for our findings. In Section 1.7.3, we assess the cost-effectiveness of the LTCI program by comparing reimbursed long-term care expenses to medical expenditures.

1.7.1 Findings on Reimbursed Formal LTC, Informal Caregiving, and Medical Expenditures

Grade 3 (Home Care Only) Benefits

Figure 1.2a displays the probability of eligibility for Grade 3 benefits (i.e. home care only) as a function of the preliminary score, and Table 1.3a the estimated increases in probability at 50 and 55. Scoring just above 50 leads to an 8 percentage point increase in the probability

[19] We conducted this exercise for all questions and responses. This example is representative of our findings.

of eligibility for home care benefits while scoring just above 55 leads to a 17 percentage point increase. To address the impact of eligibility on utilization, Figure 1.7a displays reimbursed home care expenditures as a function of the preliminary score. Note that the pattern of expenditures corresponds well with the pattern of eligibility. In particular, as the score increases from 50 to 55, home care expenditures increase with the probability of eligibility for home care benefits. Moreover, there are discrete increases in expenditures at 50 and 55 corresponding to the discrete increases in the probability of eligibility for home care benefits at those points. Panel A of Table 1.4 contains estimates of the increases in reimbursed home care expenditures at 50 and 55. The increase in eligibility at 50 leads to a \$300 increase in reimbursed home care expenditures while the increase in eligibility at 55 leads to a \$850 increase. Regarding institutional care, Figure 1.7b displays reimbursed facility care days as a function of the preliminary score and Panel B of Table 1.4 contains estimates of the corresponding increases at 50 and 55. Consistent with no change in facility care benefits, the increases in eligibility for Grade 3 benefits at 50 and 55 do not lead to a statistically significant increase in facility care use.

We now assess the corresponding impacts of these changes in reimbursed formal care utilization on informal care. Figure 1.8 displays the one year changes in the probabilities of living independently (living alone or with one's spouse) and having a child caregiver as functions of the preliminary score. Figure 1.8a shows that the probability of living independently over time falls across all scores as individuals get sicker on average. Moreover, the decrease is larger for those who were not eligible for Grade 3 benefits relative to those who were. In particular, the pattern corresponds to the pattern of reimbursed home care utilization. Despite the overall patterns, however, the increased utilization of reimbursed home care at the thresholds does not translate to a statistically significant change in the probability of living independently as estimated in Panel D of Table 1.4. We find similar results for child caregiving. As seen in Figure 1.8b, the change in child caregiving is positive across all scores as individuals age and become sicker over time. However, it increases trivially among those eligible for Grade 3 benefits, suggesting that formal home care is able to avert the use of informal care. Moreover, the use of child caregiving increases among those who were not eligible for Grade 3 benefits. Again, however, despite the overall patterns, the increased utilization at the thresholds is not associated with a statistically significant change in child caregiving as estimated in Panel C of Table 1.4.

There are several possible explanations for the limited impact on informal care. One potential explanation is that individuals who are ineligible for home care benefits may be able to finance these services privately, so that the probability of living independently (having a child caregiver) would fall (increase) less than in the absence of such an option. Another potential explanation is that formal home care allows a partial reduction, as opposed to complete elimination, of informal care. In other words, while there is no estimated impact on the extensive margin, there may still be an impact on the intensive margin. We address these potential explanations in Section 1.7.2.

Lastly, we assess the impact of increased home care utilization on medical expenditures and hospital utilization. Figure 1.9 displays the one year changes in these measures as functions of the preliminary score. We find no evidence that home care use impacts these outcomes, both across scores and treatment regimes as well as at the thresholds. The latter estimates are confirmed in Panels E and F of Table 1.4. The finding of no impact on medical expenditures is perhaps unsurprising given that the primary purpose of long-term care is not so much to restore or maintain health as it is to increase the general quality of life for the individual. We discuss these findings further in Section 1.7.3.

In summary, we find that eligibility for reimbursed home care benefits leads to the utilization of reimbursed formal home care. However, the use of reimbursed formal home care has no statistically significant impact on the use of informal care at the extensive margin nor on other medical utilization. There are various possible explanations for explaining the lack of an impact on informal care, which we address in Section 1.7.2.

Grade 2 (Home or Institutional Care) Benefits

We now assess the impact of Grade 2 benefits (i.e. where individuals can choose between home and institutional care benefits) on our outcomes of interest. Figure 1.2b displays the probability of eligibility for Grade 2 benefits as a function of the preliminary score, and Table 1.3b the estimated increases in probability at 70 and 75. Scoring just above 70 leads to a 4 percentage point increase in the probability of eligibility for home and institutional care benefits while scoring just above 75 leads to a 37 percentage point increase. To address the impact of eligibility on utilization, Figure 1.10 displays reimbursed home care expenditures and facility care days as a function of the preliminary score. We see that the pattern of reimbursed institutional care days corresponds well with the pattern of eligibility for those benefits. Consequently, reimbursed home care expenditures decrease as individuals substitute facility care for home care. Moreover, there are discrete increases (decreases) in facility (home) care use corresponding to the discrete increases in the probability of eligibility for institutional care at 70 and 75. Panels A and B of Table 1.4 contains estimates of the increases in reimbursed formal care expenditures at 70 and 75. The increase in eligibility at 70 leads to a 24 day increase in reimbursed facility use and a $400 decrease in home care expenditures. The increase in eligibility at 75 leads to a 23 day increase in reimbursed facility use and a $550 decrease in home care expenditures.

We next assess corresponding changes in informal care. Figure 1.11 displays the one year change in the probabilities of living independently and having a child caregiver as functions of the preliminary score. Again, we see that the change in the probability of living independently is negative across all scores as individuals get sicker over time, with the reduction slightly stronger for individuals eligible for facility benefits. However, there is no statistically significant change in independent living corresponding to the change in long term care utilization at 70 and 75 as estimated in Panels D of Table 1.4. For child caregiving, we see that it falls with the onset of facility care benefits, mimicking the pattern of eligibility for Grade 2 benefits. There is also suggestive evidence that the increased utilization of facility care benefits over home care benefits at 70 translates to a reduction in child caregiving, consistent with estimates in Panel C of Table 1.4. Estimates at our preferred bandwidth suggest that Grade 2 benefits lead to a statistically significant decrease in the probability of child caregiving of 3 percentage points. Estimates at more stringent bandwidths, including the IK, suggest similarly negative impacts, but these estimates are not precise enough to be statistically significant. Similarly for 75, estimates suggest negative, but not statistically significant, impacts on child caregiving.

There are several possible explanations for these findings. That there is no impact on independent living may not be a surprise. While facility care substitutes for home care, they both are linked to dependent living situations. Although we do not find impacts of home care on the use of child caregiving, we do find suggestive impacts of facility care on the use of child caregiving. This is consistent with the fact that formal home care may reduce but not completely eliminate child caregiving. It is less likely that significant child caregiving would continue while the care recipient resides in a facility. We address these considerations more carefully in Section 1.7.2.

Lastly, we look at the impact of increased facility care and decreased home care utilization on medical expenditures and hospital utilization. Figure 1.12 displays the one year changes in these measures as functions of the preliminary score. We find no evidence that the substitution of facility care for home care at 70 impacts these outcomes. However, there is suggestive evidence at 75 that the substitution of facility care for home care leads to reductions in medical expenses and that this is largely accounted for by a reduction in hospital expenses. The estimates are shown in Panels E and F of Table 1.4. One explanation for this finding is that these individuals in this setting are less likely to experience costly accidents. Another explanation is that patients are able to transition sooner out of more expensive hospital care and into less expensive facility care. We discuss these findings further in Section 1.7.3.

In summary, we find that eligibility for facility care benefits leads to the substitution of facility care for home care. There is no impact on independent living, but there is suggestive evidence that this leads to a reduction in child caregiving at the extensive margin. There is also evidence for a corresponding reduction in medical utilization. As in our analysis of Grade 3 benefits, it will be important to take into account the ability of individuals to pay for formal long-term care services out of pocket, which we address in Section 1.7.2.

Grade 1 (Increased Maximum for Home Care, Increased Price for Institutional Care) Benefits

We now assess the impact of Grade 1 benefits on our outcomes of interest. Recall that these benefits are effectively an increase in the maximum benefit for home care combined with a discontinuous increase in the cost of facility care at the threshold. Figure 1.2c displays the probability of eligibility for Grade 1 benefits as a function of the preliminary score, and Table 1.3c the estimated increases in probability at 90 and 95. Scoring just above 90 does not lead to a statistically significant increase in eligibility for Grade 1 benefits. Thus, assessments at this threshold serve as placebo tests for this design. As expected, we find no statistically significant impacts on reimbursed home expenditures and facility days, child caregiving and living independently, and medical and hospital expenses (see Figures 1.13 to 1.15 and the fifth row of Table 1.4).

A preliminary score just above 95 leads to an 83 percentage point increase in the probability of eligibility for Grade 1 benefits. To address the impact of eligibility on utilization, Figure 1.13 displays reimbursed home care expenditures and facility care days as functions of the preliminary score, and Panels A and B of Table 1.4 corresponding estimates of the discontinuities. Due to how Grade 1 benefits lead to a relative price increase in facility care, Grade 1 benefits at 95 lead to a 30 day decrease in the number of facility days used and a $930 increase in reimbursed home expenditures. As shown in Figure 1.14, with corresponding estimates in Panels C and D of Table 1.4, this shift in formal long-term care mix is not statistically significantly associated with changes in informal care, as measured by child caregiving and independent living. However, as shown in Figure 1.15 and Panels E and F of Table 1.4 we do find a statistically significant decrease in medical expenses of almost $700, coupled with a decrease in hospital expenditures of nearly the same amount. The fact that we find an impact of home care on medical expenditures in this case but not for Grade 3 may be due to the fact that individuals who receive Grade 1 benefits are more frail and susceptible to health shocks that can be ameliorated by formal care. We discuss our findings on medical expenditures further in Section 1.7.3.

In summary, we find that a relative increase in the price of facility care leads to increased utilization of formal home care. This shift in formal long-term care services has no impact on informal care but has a substantial impact on medical expenses, largely due to decreased hospital expenditures.

1.7.2 Crowd Out and Informal Care Intensity

The analysis of Grade 3 benefits in Section 1.7.1 indicated that an increase in reimbursed home care expenditures had little impact on informal care as measured by independent living and child caregiving. One possible explanation for this finding is that public financing simply crowds out private expenditures for home care. Another possible explanation is that publicly financed home care enables individuals to reduce informal caregiving at the intensive margin. Unfortunately, our data does not provide measures of private spending on home care, nor does it contain measures of the amount of caregiving. Instead we focus on a subpopulation of individuals—those in the MCA program and thus are poor—for whom the likelihood of out-of-pocket spending is expected to be low.

The first column of Table 1.5 indicates estimates of the increase in home care utilization at 50 and 55 for the subset of MCA individuals. As in the overall sample, Grade 3 benefits lead to an increase in home care expenditures at 50 and 55 for MCA individuals. Columns two and three indicate estimates of the change in informal care at 50 and 55. As in the overall population, there is no statistically significant impact of Grade 3 benefits on informal care at the extensive margin for MCA individuals. Given that MCA individuals are unlikely to pay for home care out of pocket, these results suggest that the lack of an observable impact on informal care is not likely to be solely due to crowd out of private spending on formal care by public reimbursement.

A remaining explanation for why public reimbursement has no impact on informal care at the extensive margin is that the impact is on the intensive margin. To shed light on this possibility, we look at the impact of Grade 3 benefits on the use of a particular home care service, short-term respite care. Short-term respite care is short-term (i.e. a few days) facility care used to provide temporary relief for the regular caregiver. Thus, use of this type of home care is a strong indication for reduction in informal caregiving at the intensive margin. Indeed, as shown in Table 1.6, which shows estimates for several home care services, we find Grade 3 benefits lead to a statistically significant increase in the use of short-term respite care at 55.

As in the case for home care, we only observe publicly financed facility care. To measure the extent of crowd out we need a measure of all facility care, regardless of whether it is financed publicly or privately. To accomplish this we use an indirect measure of all facility utilization: medical spending occurring in a long-term care facility (i.e. regardless of financing). If the probability of having medical spending occurring in a long-term care facility is a fixed percentage of those who attend a long-term care facility (at the threshold) then changes in the probability of having medical spending occurring in a long-term care facility will capture changes in the probability of attending a long-term care facility. In other words, if $\frac{\text{\# w/Medical Spending in LTC Facility}}{\text{\# in LTC Facility}}$ is fixed, then a percentage increase in the denominator will be tied to a percentage increase in the numerator of the same magnitude.[20] Table 1.7 presents estimates of the impact at 70 and 75 of the probability of using a publicly financed long-term care facility and the probability of having medical spending occurring in a long-term care facility. Scoring just above 70 is associated with a 25% increase (6.5 percentage points on a base of 25.7%) in the probability of using publicly financed facility care. However, using the probability of medical spending occurring in a long-term care facility as a proxy for all facility care shows that the probability of using facility care, regardless of financing, increases only 18.4% (2.9 percentage points on a base of 15.7%) at 70. This suggests that about a quarter of publicly financed care is used to substitute for out of pocket expenditures. The corresponding measure of crowd out at 75 is 46.7%. The fact that crowd out is higher at 75 than 70 is not surprising, given that individuals at 75 have more need for long-term care and thus are more likely to privately finance facility care in the absence of LTCI. While these measures of crowd out are substantial, they also suggest that crowd out is not complete, and therefore cannot fully explain our lack of findings for informal care.

1.7.3 LTC Expenditures and Reductions in Medical Expenses

In light of the previous results showing decreases in medical expenditure, a useful metric for assessing the cost-effectiveness of this policy and its costs to the government is to compare the reimbursed long-term care expenses to the changes in medical expenses. Recall that with the administrative data we use, we are able to measure the both the universe of medical

[20]It is possible that those who spend out of pocket (i.e. those below the threshold) are likely to be sicker and thus have a higher probability of medical spending occurring in a facility. To the extent that this is the case, we will find a smaller change in the probability of having medical spending occurring in a facility and an over (upper bound) estimate of crowdout.

expenditures and the universe of reimbursed long-term care expenditures. The first set of columns of Table 1.9 display the estimated impacts of all thresholds on reimbursed long-term care expenditures. For convenience, the second set of columns redisplay the impacts on medical expenditures. The third set of columns indicate the medical expenditures saved per additional dollar of long-term care expenditure reimbursed.

A preliminary score above 50 and 55 leads to a \$208 and \$931 increase in total reimbursed long-term care expenditures, respectively. As seen earlier, however, this results in little, if any, savings in medical expenditures. Focusing on Grade 2, we see that additional benefits for facility care lead to an additional \$500 in expenditures as individuals substitute more expensive facility care in place of home care. However, corresponding to this increase in expenditure we find a decrease in medical expenditures of more than \$300, for a medical expenditure savings of \$0.6 per dollar of long-term care reimbursed. The fact that there is no apparent savings at 70 may be due to heterogeneous impacts of the policy or possible bias at 75. Focusing on Grade 1 at 95 (recall that there is little effective change in eligibility at 90), we see that additional benefits for Grade 1 lead to only small changes in expenditures as individuals tend to use more home care and less facility care. However, this substitution leads to large impacts on medical expenditures—nearly a \$700 reduction. Thus, Grade 1 benefits lead to a medical expenditures savings of more than \$650 per dollar of long-term care reimbursed. Clearly, the amount of long-term care reimbursed is not a complete measure of the costs of the program as it does not include the administrative expenses, for example. Moreover, medical expenses are not a complete measure of the potential cost savings of the program as impacts on labor outcomes could have impacts on government revenue.[21] However, the large impact we measure here highlights the importance of considering the potential program savings from reduced medical expenditures.

[21] Our limited findings on informal care at the extensive margin suggest that these labor market impacts are likely small.

1.8 Robustness

1.8.1 Balance of Covariates

As discussed in Section 1.3.2, an important assumption for our identification strategy is that individuals on each side of the thresholds are comparable. A test of this assumption is to check the balance of observable characteristics across the thresholds. Table 1.10 contains estimates of the discontinuities around the thresholds for predetermined variables that are likely to be correlated with our dependent variables of interest. With the exception of the 75 threshold, most of the variables appear to be continuous around the thresholds at our preferred bandwidth.

Because we are testing numerous variables and thresholds, some discontinuities will be statistically significant by random chance. As a result, we conduct two tests which account for this, with results presented in the last columns of Table 1.10. First, we look at a summary measure—the predicted medical expenditures from a regression of medical expenditures on the other predetermined variables. Again, with the exception of the 75 threshold, there appear to be no discontinuities in predicted medical expenditures at our preferred bandwidth. Second, we test whether the discontinuities are jointly significant by seemingly unrelated regression, as described in Lee and Lemieux (2010). Consistent with the first exercise, the only threshold for which the discontinuities are jointly significant at the preferred bandwidth is 75. This leads us to believe that our results are not impacted by unobserved confounders at the other thresholds. Nonetheless, we controlled for the few instances of significance occurring in our variables of interest by estimating differences in our dependent variables in our regressions.

1.8.2 Differential Mortality

Another relevant outcome is whether these benefits had any impact on mortality. This measure is important in and of itself, and is useful because it is objective and well-defined. Moreover, it is important to address the concern that differential mortality around the thresholds could account for our findings. For example, if individuals just below the threshold were more likely to die as a result of not receiving treatment, relatively healthy individuals would remain in the sample, minimizing any estimated impacts. We assess this by looking at mortality by 2010 around the thresholds. Table 1.8 displays estimates of Equation 1.1 with mortality by 2010 as the outcome. We find no statistically significant differences in mortality at all thresholds. Thus, the increase in long-term care utilization at the thresholds has no impact on mortality in the short-run.

1.8.3 Other Specifications

A consequential decision in estimating Equation 1.1 is the choice of bandwidth. Although we have shown that our results are qualitatively consistent at both our preferred bandwidth and the IK bandwidth, it is useful to know how sensitive our findings are to bandwidth choice. To do so, we reestimate Equation 1.1 for our main outcomes of interest at several bandwidths—from 1 to 5, in increments of 0.5. Figures B.3 to B.8 plot the estimated coefficients with 95% confidence bands against the bandwidth. There are two things worth highlighting. First, coefficients are less precisely estimated and more variable at very small bandwidths. Second, the coefficient estimate at our preferred bandwidth falls within the 95% confidence bands of the estimates at other bandwidths in general, indicating that our findings are not too sensitive to bandwidth selection.

On the specification of $f(S)$, our approach in this paper follows Hahn, Todd, and van der Klaauw (2001) by using local linear regressions to estimate the discontinuity at the threshold. As shown in the previous section, our findings are consistent even at very small bandwidths. Moreover, visual inspection suggests the relationship between eligibility (as well as our outcomes of interest) and the preliminary score is fairly linear even at relatively large distances from the thresholds. Nonetheless, in Figures B.9 to B.14 we explore how sensitive our findings are to higher order specifications of $f(S)$ at our preferred bandwidth. For the most part, the coefficient estimate based on a linear specification of $f(S)$ falls within the 95% confidence bands of estimates for higher order specifications. However, the variance of the higher order specifications grows quite large, which lends support for the use of linear splines.

1.8.4 Differences-in-Differences Estimation

Our research design takes advantage of a setting with a continuous measure of long-term care needs (i.e. the preliminary score) and thresholds that lead to "as good as random" variation in the probabilities of benefits. One limitation of this design, however, is the reduced precision from relying primarily on observations around the threshold. In this section, we estimate a differences-in-differences model that relies on stronger assumptions, but has potentially improved precision. Specifically, we compare three groups of individuals: individuals who are treated based solely on the preliminary score (for Grade 3, we consider individuals with preliminary scores in [55,60)), individuals who are treated based on committee guidelines (for Grade 3, these are individuals with preliminary scores in [50,55)), and individuals who are not treated (for Grade 3, these are individuals with preliminary scores in [45,50)). For $\tau \in \{55, 75, 95\}$, we define $\text{commit}_\tau \equiv 1\!\!1\{\tau - 5 \leq S < \tau\}$ and $\text{treat}_\tau \equiv 1\!\!1\{\tau \leq S < \tau + 5\}$, where S is the 2009 preliminary score. With the untreated individuals (i.e. $\{S : \tau - 10 \leq S < \tau - 5\}$) as our reference group, we estimate the following saturated differences-in-differences model for an individual i in a one point bin b at time t:

$$\text{outcome}_{ibt} = \sum_{t \neq 0} (\beta_t^C \text{commit}_\tau \cdot \phi_t + \beta_t^T \text{treat}_\tau \cdot \phi_t) + \gamma_i + \phi_t + \eta_b \cdot t + \epsilon_{ibt}, \tag{1.2}$$

where γ_i is an individual-specific effect, ϕ_t a year-fixed effect, $\eta_b \cdot t$ is a bin-specific linear time trend, and the baseline year is set to $t = 0$.[22] The key assumption underlying this estimation method is that there are no unobserved factors that affect the three groups differentially within a year.

Table 1.11 presents estimates of β_1^C and β_1^T from Equation 1.2. Grade 3 expenditures lead to a statistically significant decrease in child caregiving, but have no statistically significant impact on independent living. There is no statistically significant impact on medical expenditures or hospital expenses. Additional long-term care expenditures resulting from Grade 2 benefits are also associated with a statistically significant decrease in child caregiving, but not independent living. The use of Grade 2 benefits leads to a decrease in other medical expenses, accounted for largely by hospital expenses. These results translate into a medical dollars saved per additional dollar of reimbursed long-term care expenditure of 0.2–0.3. Grade 1 benefits are associated with a statistically significant increase in child caregiving, but not independent living. This is consistent with the increased use of home care among these individuals that was found earlier. The use of Grade 1 benefits leads to a decrease in other medical expenses, largely accounted for by hospital expenses. In this case, the medical dollars saved per additional dollar of reimbursed long-term care expenditure is more than one, suggesting strong program savings.

The findings from this analysis are fairly consistent with our findings from the regression discontinuity analysis. Even though the differences-in-differences analysis suggests statisti-

[22] For our outcome measures (medical expenditures and informal care), $t \in \{-1, 0, 1\}$. For our long-term care utilization measures, $t \in \{0, 1, 2\}$. Table 1.11 presents estimates of β_1^C and β_1^T from Equation 1.2.

cally significant impacts on child caregiving while RD estimates do not, this could be due to lack of statistical precision. Moreover, estimates of medical expenditures saved per dollar of reimbursed long-term care are similar across both estimation strategies.

Lastly, this estimation strategy allows us to compare the committee affected group to the automatically treated group. This is particularly relevant given that assigning treatment based solely on the preliminary score may not be optimal and that leaving room for discretionary assignment of treatment may improve efficiency. In this analysis, there do not appear to be any striking differences in performance between the two groups among Grades 3 and 2 individuals. However, it appears that the committee affected group has a more substantial impact among Grade 1 affected individuals. While this suggests the possibility that a more discretionary decision-making procedure for determining treatment may be more effective than a hard rules-based criteria, we caution that this measure (vs. quality of life, for example) may not be the primary objective to optimize from the standpoint of the committee.

1.9 Discussion

In this paper, we find that publicly financed LTCI leads to small, if any, impacts on informal care at the extensive margin. We determine that this is not solely due to crowdout, but partly explained by the fact that informal care is reduced at the intensive margin. That we find limited impacts on informal care stands in contrast to some of the previous literature, but is not surprising given that South Korea is a strong family ties country. That is, due to family obligations, Koreans may find it more difficult to give up completely the responsibility of taking care of their elderly parents. That we still find reductions in the intensive margin indicate that our results constitute a lower bound for the effect in the U.S. in general, and may be directly indicative of population subgroups in the U.S. such as Asians and Hispanics.

Interestingly, we find that among people who are partially dependent for several activities of daily living, transitioning from home to facility care results in decreased medical expenditures. This may come as a surprise at first, given that the purpose of long-term care is not so much to restore or maintain health as it is to increase the general welfare of the individual by facilitating activities of daily living. Indeed, we find no impacts on health as captured by mortality. However, a plausible explanation is that the increased attention one receives in a facility may prevent costly accidents like falling and breaking one's hip. Another possibility is that patients are able to transition sooner out of more expensive hospital care and into less expensive facility care. Surprisingly, among individuals who are completely dependent for several activities of daily living, the opposite transition leads to substantially lower medical expenses. This may be mediated by the fact that the presence of medical professionals in a facility may lead to additional or more costly care than if one were being cared for at the home, and that, among this population of individuals, this effect predominates the previously mentioned effects. In fact, that transitioning people from institutions to the community may be beneficial is consistent with the objectives of programs such as Money Follows the Person in the U.S. This supports the more general point that our findings on medical expenses are not culturally or context specific, and that understanding the relationship between long-term care expenses and medical expenses may be a fruitful avenue to contain health care costs.

1.10 Conclusion

Results from this paper provide insight into the welfare impacts of government reimbursement of long-term care on care recipients, caregivers, and taxpayers, as well as suggestions for the design of optimal long-term care policy. Our main finding is that the benefits of home and facility care are heterogeneous across physical function level and therefore setting policy accordingly has the potential to dramatically reduce medical expenses. We also find that formal long-term care is not a strong substitute for informal long-term care at the extensive margin.

Among individuals who are partially dependent for some activities of daily living we find that government subsidies for formal home care lead to an overall increase in its utilization, even accounting for crowd out, with no impact on informal caregiving at the extensive margin, medical expenses, or mortality. While we find evidence for a reduction in informal caregiving at the intensive margin, this suggests that if the policy objective is to increase the labor supply of individuals caring for this population, subsidies for home care may have limited impact. Moreover, the converse of our findings on medical expenses and mortality suggest that home care reimbursement may be reduced without significant detriment to the health of the care recipient.

Among individuals who are partially dependent for several activities of daily living, additional reimbursement of institutional care leads to the crowd out of privately financed institutional care of up to 47%. Institutional care does increase overall, leading to reductions in informal caregiving and medical expenses. From a policy perspective, the latter finding suggests that while substitution of institutional care for less expensive home care may lead to increased costs, this may be partially offset by reductions in medical expenses. Moreover, our finding on informal caregiving suggests that this policy may lead to increased labor supply of individuals caring for this population. In this case, optimal policy depends on the objective function of the policymaker in balancing the tradeoff between increased taxpayer costs, reduced informal caregiving, and improved quality of life for the care recipient.

Among individuals who are completely dependent for several activities of daily living, we find that an increase in the price of institutional care combined with an increase in the benefit maximum for home care leads to substitution of home care for institutional care. While we find no impact on informal caregiving, we find substantial decreases in medical spending. From a policy perspective, this suggests that increased incentives for the use of home care may lead to an improvement in the welfare of care recipients while limiting or even reducing costs to taxpayers.

Table 1.1: Overview of Grades of Benefits

Classification	Description	Criteria	Home Care Max Benefit (USD)	Institutional Care Max Benefit (USD)
No Benefits		score < 55	none	none
Grade 3	Need assistance moving around, partially dependent for some ADLs	55 ≤ score < 75	750 / month	none
Grade 2	Unable to move on own, partially dependent for several ADLs	75 ≤ score < 95	900 / month	40 / day
Grade 1	Bedridden, completely dependent for several ADLs	95 ≤ score	1100 / month	45 / day

1 USD ≈ 1100 KRW

Table 1.2: Summary Statistics by Grade

Adjusted Score	No Benefits [45,55)	Grade 3 [55,75)	Grade 2 [75,95)	Grade 1 [95+]
# Obs	*35,580*	*43,615*	*76,170*	*12,090*
Age	76.06	76.87	78.13	77.01
	(8.16)	(8.85)	(8.74)	(9.68)
Female	0.77	0.73	0.74	0.73
	(0.42)	(0.44)	(0.44)	(0.44)
Urban	0.73	0.76	0.78	0.77
	(0.45)	(0.43)	(0.41)	(0.42)
Insurance Contribution	41.27	54.32	62.49	64.19
	(63.92)	(73.04)	(74.51)	(79.54)
MCA	0.43	0.31	0.23	0.25
	(0.50)	(0.46)	(0.42)	(0.43)
ADL Index	17.42	20.07	24.96	30.05
	(3.78)	(4.48)	(5.63)	(5.99)
Medical Expenditures	2,255	2,850	4,165	5,080
	(4,312)	(5,190)	(6,719)	(8,060)
Hospital Days	12.53	19.56	42.48	55.37
	(42.53)	(57.86)	(89.82)	(105.59)
Child Caregiver	0.26	0.30	0.23	0.19
	(0.44)	(0.46)	(0.42)	(0.39)
Live Independently	0.60	0.42	0.21	0.21
	(0.49)	(0.49)	(0.41)	(0.40)
LTC Facility Days	21.90	66.82	159.47	158.58
	(63.69)	(137.56)	(167.55)	(168.94)
Home Care Exp	2,885	5,061	3,442	3,384
	(3,037)	(3,836)	(4,263)	(4,682)

Notes: Sample consists of individuals who were assessed for long-term care insurance in 2008 and 2009. Grade categorization is based on the 2009 adjusted score. All measures are at baseline, except for long-term care facility days and home care expenditures. See text for definitions of variables.

Table 1.3: Effect of Thresholds on Changes in Eligibility

(a)

Grade 3 Eligibility			
Bandwidth	**2.5**	**IK**	**IK B/W**
Score $\geq$ 50	0.08**	0.09**	1.3
	(0.01)	(0.02)	
Score $\geq$ 55	0.17**	0.10**	0.6
	(0.01)	(0.03)	

(b)

Grade 2 Eligibility			
Bandwidth	**2.5**	**IK**	**IK B/W**
Score $\geq$ 70	0.04**	0.06**	0.8
	(0.01)	(0.01)	
Score $\geq$ 75	0.37**	0.39**	1.1
	(0.01)	(0.03)	

(c)

Grade 1 Eligibility			
Bandwidth	**2.5**	**IK**	**IK B/W**
Score $\geq$ 90	0.01+	0.01*	0.9
	(0.00)	(0.00)	
Score $\geq$ 95	0.83**	0.83**	1.1
	(0.01)	(0.02)	

Notes: The first two columns of each panel report estimates of β from local linear regression of Equation (1.1). Each cell represents a different regression. The running variable is the 2009 preliminary score. Controls include age, gender, region type, insurance type, and insurance contribution. Rectangular kernel. The third column of each panel reports the optimal bandwidth determined by the IK procedure. Robust standard errors in parentheses. ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

Table 1.4: Main Results on LTC Utilization, Informal Care, and Medical Expenditures

	Reimbursed Formal LTC Utilization						Informal Care						Medical Utilization					
	Panel A			Panel B			Panel C			Panel D			Panel E			Panel F		
	LTC Home Expenditures			LTC Facility Days			Pr(Child Caregiver)			Pr(Live Independently)			Medical Expenses			Hospital Expenses		
Bandwidth	**2.5**	**IK**	**IK B/W**	**2.5**	**IK**	**IK B/W**	**2.5**	**IK**	**IK B/W**	**2.5**	**IK**	**IK B/W**	**2.5**	**IK**	**IK B/W**	**2.5**	**IK**	**IK B/W**
Score ≥ 50	311*	480**	2.0	-2.3	-4.5	1.0	-0.03	0.13*	1.4	-0.02	-0.14**	1.1	-97	-198	1.8	-177	-241	1.9
	(157)	(181)		(3.7)	(6.5)		(0.04)	(0.06)		(0.04)	(0.05)		(174)	(217)		(174)	(213)	
Score ≥ 55	850**	748**	1.7	0.2	24.2**	1.0	0.01	0.00	1.1	-0.02	-0.02	1.3	59	-67	1.9	60	2	3.1
	(134)	(163)		(4.0)	(6.6)		(0.02)	(0.03)		(0.02)	(0.02)		(146)	(171)		(141)	(127)	
Score ≥ 70	-392**	-561**	1.3	23.8**	37.8**	1.2	-0.03*	-0.02	1.2	0.00	0.00	1.2	101	86	2.6	145	93	2.7
	(145)	(208)		(5.3)	(8.2)		(0.01)	(0.02)		(0.02)	(0.02)		(173)	(170)		(176)	(170)	
Score ≥ 75	-554**	-571**	2.1	22.5**	24.7**	1.0	-0.01	-0.03	1.2	-0.02	0.00	1.1	-327+	-405*	2.1	-370*	-125	1.5
	(141)	(155)		(5.4)	(9.1)		(0.01)	(0.02)		(0.02)	(0.02)		(178)	(200)		(183)	(242)	
Score ≥ 90	25	-21	2.2	3.7	2.6	2.5	0.02	-0.01	1.3	0.02	0.02	1.4	-344	-249	2.7	-245	-149	2.9
	(219)	(231)		(8.5)	(8.4)		(0.02)	(0.03)		(0.02)	(0.02)		(295)	(283)		(302)	(289)	
Score ≥ 95	926**	586+	1.6	-29.5**	-27.9*	1.5	0.02	0.04	1.5	0.00	0.00	1.5	-691*	-757*	2.3	-666+	-727*	2.3
	(242)	(320)		(9.0)	(11.9)		(0.02)	(0.03)		(0.03)	(0.03)		(319)	(333)		(342)	(356)	

Notes: The first two columns of each panel in this table report estimates of β from local linear regression of Equation (1.1). Each cell represents a different regression. The running variable is the 2009 preliminary score. Controls include age, gender, region type, insurance type, and insurance contribution. Rectangular kernel. The third column of each panel reports the optimal bandwidth determined by the IK procedure. Robust standard errors in parentheses. ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

Table 1.5: Utilization and Informal Care for MCA Individuals

	LTC Home Expenditures			Pr(Child Caregiver)			Pr(Live Independently)		
Bandwidth	2.5	IK	IK B/W	2.5	IK	IK B/W	2.5	IK	IK B/W
Score $\geq$ 50	476+	577	1.5	0.06	0.05	1.9	-0.06	-0.06	1.9
	(247)	(355)		(0.06)	(0.06)		(0.06)	(0.06)	
Score $\geq$ 55	930**	-358	1.0	0.03	0.04	1.6	-0.04	-0.04	1.6
	(232)	(371)		(0.02)	(0.03)		(0.03)	(0.03)	

Notes: The first two columns of each panel in this table report estimates of β from local linear regression of Equation (1.1). Each cell represents a different regression. The running variable is the 2009 preliminary score. Controls include age, gender, region type, insurance type, and insurance contribution. Rectangular kernel. The third column of each panel reports the optimal bandwidth determined by the IK procedure. Sample consists of individuals in the MCA program. Robust standard errors in parentheses. ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

Table 1.6: Detailed Home Care Utilization

	Home Help		Home Bath		Home Nursing		Day / Evening Care		Respite Care		Equipment	
Bandwidth	2.5	IK	2.5	IK	2.5	IK	2.5	IK	2.5	IK	2.5	IK
Score $\geq$ 50	11.42*	11.44	0.04	0.15	0.06	-0.06	-2.43	-0.56	-0.74	-1.76	2.85	1.93
	(4.76)	(7.44)	(0.45)	(0.45)	(0.20)	(0.25)	(1.81)	(1.67)	(1.31)	(1.45)	(2.07)	(1.97)
Score $\geq$ 55	16.02**	1.28	0.50	0.20	-0.33	-0.39	2.39	3.49	6.67**	6.76**	1.25	0.05
	(4.21)	(5.73)	(0.42)	(0.55)	(0.25)	(0.24)	(2.11)	(2.49)	(1.38)	(1.38)	(2.22)	(2.53)

Notes: Each cell reports estimates of β from a different local linear regression of Equation (1.1). Dependent variables are measured in # of visits. The running variable is the 2009 preliminary score. Rectangular kernel. Optimal bandwidths for the IK procedure are omitted for space considerations and are available from the authors upon request. Robust standard errors in parentheses. ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

Table 1.7: Crowd Out of Facility Care

Bandwidth	Pr(Publicly Financed Facility Care) 2.5	Pr(Med Spending in LTC Facility) 2.5	Crowd Out (%)
Change at 70	0.065** (0.016)	0.029* (0.0135)	
Base at 70	0.257** (0.011)	0.156** (0.009)	
% Change at 70	25.4%	18.4%	**27.4%**
Change at 75	0.068** (0.017)	0.016 (0.014)	
Base at 75	0.483** (0.015	0.214** (0.012)	
% Change at 75	14.2%	7.5%	**46.7%**

Notes: Columns 1 and 2 report coefficient estimates from Equation (1.1). Dependent variables are indicators for public reimbursement of facility care and medical spending in a LTC facility. The running variable is the 2009 preliminary score. Rectangular kernel. "Change at 'X' " is the estimate of β. "Base at 'X' " is the predicted value of the dependent variable at 'X' minus the "Change at 'X' ". ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

Table 1.8: Effect of Eligibility on Mortality

	Mortality by 2010		
Bandwidth	2.5	IK	IK B/W
Score ≥ 50	0.00	-0.02	1.1
	(0.01)	(0.02)	
Score ≥ 55	0.00	0.00	1.2
	(0.01)	(0.01)	
Score ≥ 70	0.00	0.00	1.2
	(0.01)	(0.02)	
Score ≥ 75	0.00	0.02	0.9
	(0.01)	(0.02)	
Score ≥ 90	0.00	0.03	1.5
	(0.02)	(0.03)	
Score ≥ 95	-0.02	-0.04	1.5
	(0.02)	(0.03)	

Notes: The first two columns of this table report estimates of β from local linear regression of Equation (1.1). Each cell represents a different regression. The dependent variable is mortality by 2010. The running variable is the 2009 preliminary score. Controls include age, gender, region type, insurance type, and insurance contribution. Rectangular kernel. The third column of each panel reports the optimal bandwidth determined by the IK procedure. Robust standard errors in parentheses. ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

Table 1.9: LTC Expenses vs. Medical Care Savings

	LTC Expenditures			Medical Expenses			$ Med Exp Saved / $ LTC Spent	
Bandwidth	**2.5**	**IK**	**IK B/W**	**2.5**	**IK**	**IK B/W**	**2.5**	**IK**
Score ≥ 50	208	32	1.8	-97	-198	1.8	0.5	6.1
	(169)	(209)		(174)	(217)			
Score ≥ 55	931**	1,090**	1.8	59	-67	1.9	-0.1	0.1
	(140)	(172)		(146)	(171)			
Score ≥ 70	524**	796**	1.9	101	86	2.6	-0.2	-0.1
	(156)	(183)		(173)	(170)			
Score ≥ 75	535**	711**	1.1	-327+	-405*	2.1	0.6	0.6
	(164)	(267)		(178)	(200)			
Score ≥ 90	155	10	2.4	-344	-249	2.7	2.2	23.9
	(259)	(268)		(295)	(283)			
Score ≥ 95	1	-379	1.2	-691*	-757*	2.3	668.3	-2.0
	(281)	(432)		(319)	(333)			

Notes: The first two columns of the first two panels in this table report estimates of β from local linear regression of Equation (1.1). Each cell represents a different regression. The running variable is the 2009 preliminary score. Controls include age, gender, region type, insurance type, and insurance contribution. Rectangular kernel. The third columns report the optimal bandwidth determined by the IK procedure. Robust standard errors in parentheses. ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$. The third panel equals the coefficient in the second panel divided by the coefficient in the first panel.

Table 1.10: Covariate Balance

	Med Exp		Med Exp in LTC		MCA Insurance		Insurance Contrib.		Child Caregiver		Live Independently		SUR	Predicted Med Exp	
Bandwidth	**2.5**	**IK**	**2.5**	**IK**	**2.5**	**IK**	**2.5**	**IK**	**2.5**	**IK**	**2.5**	**IK**	**p-value**	**2.5**	**IK**
Score ≥ 50	311	216	17	18	-0.03	-0.02	3.16	3.25	-0.04	-0.13*	0.05	0.22**	0.49	141	83
	(202)	(189)	(81)	(86)	(0.02)	(0.03)	(3.50)	(3.46)	(0.04)	(0.06)	(0.05)	(0.07)		(127)	(135)
Score ≥ 55	42	115	43	195**	-0.01	0.02	-1.17	-3.78	-0.03	-0.03	0.07**	-0.01	0.19	115	125
	(167)	(148)	(58)	(70)	(0.02)	(0.02)	(3.21)	(3.65)	(0.02)	(0.03)	(0.03)	(0.04)		(95)	(76)
Score ≥ 70	-28	-243	120	261+	-0.02	-0.05*	-3.94	-4.45+	0.00	0.01	-0.02	-0.02	0.34	183	340*
	(206)	(229)	(111)	(135)	(0.02)	(0.02)	(2.77)	(2.51)	(0.02)	(0.03)	(0.02)	(0.03)		(130)	(158)
Score ≥ 75	1,095**	591*	753**	274	-0.03+	-0.03	3.68	3.55	0.01	0.04	-0.01	-0.01	0.00	826**	509**
	(217)	(255)	(125)	(175)	(0.01)	(0.03)	(2.45)	(2.63)	(0.02)	(0.03)	(0.02)	(0.03)		(141)	(167)
Score ≥ 90	258	338	174	180	0.03	0.02	1.29	2.73	-0.02	-0.03	-0.07*	-0.11**	0.35	137	191
	(333)	(386)	(220)	(222)	(0.02)	(0.03)	(4.42)	(4.98)	(0.03)	(0.04)	(0.03)	(0.04)		(239)	(254)
Score ≥ 95	801*	410	297	303	-0.01	-0.03	6.65+	9.33+	-0.03	0.11*	0.03	0.05	0.23	465+	454+
	(396)	(330)	(249)	(285)	(0.02)	(0.04)	(3.91)	(4.88)	(0.03)	(0.05)	(0.03)	(0.04)		(274)	(275)

Notes: Columns 1-6 and 8 report estimates of β from local linear regression of Equation (1.1). Each cell represents a different regression. Dependent variables are 2008 measures. The running variable is the 2009 preliminary score. Rectangular kernel. Optimal bandwidths for the IK procedure are omitted for space considerations and are available from the authors upon request. Robust standard errors in parentheses. ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$. Column 7 reports the p-value from a joint test of the coefficients in each row from a SUR where the bandwidth is 2.5.

Table 1.11: Differences-in-Differences Estimates

Grade 3	LTC Exp	Med Exp	Hosp Exp	Child Caregiver	Live Independently	$ Med Exp Saved / $ LTC Spent
Committee	982.1**	41.40	105.5	-0.0227	-0.0111	0.0
	(62.85)	(92.88)	(88.27)	(0.0162)	(0.0168)	
Treatment	3,141**	-11.50	22.99	-0.0604**	0.0182	0.0
	(51.97)	(79.67)	(74.13)	(0.0133)	(0.0139)	
Individuals	33,005	33,005	33,005	32,997	32,061	
R-squared	0.571	0.006	0.004	0.003	0.021	

Post-period regression coefficients from differences-in-differences estimation with individual fixed effects, year fixed effects, and bin-specific linear time trends. Committee = Preliminary Score in [50,55), Treatment = Preliminary Score in [55,60), Omitted = Preliminary Score in [45,50). Robust standard errors in parentheses.
** p<0.01, * p<0.05, + p<0.1

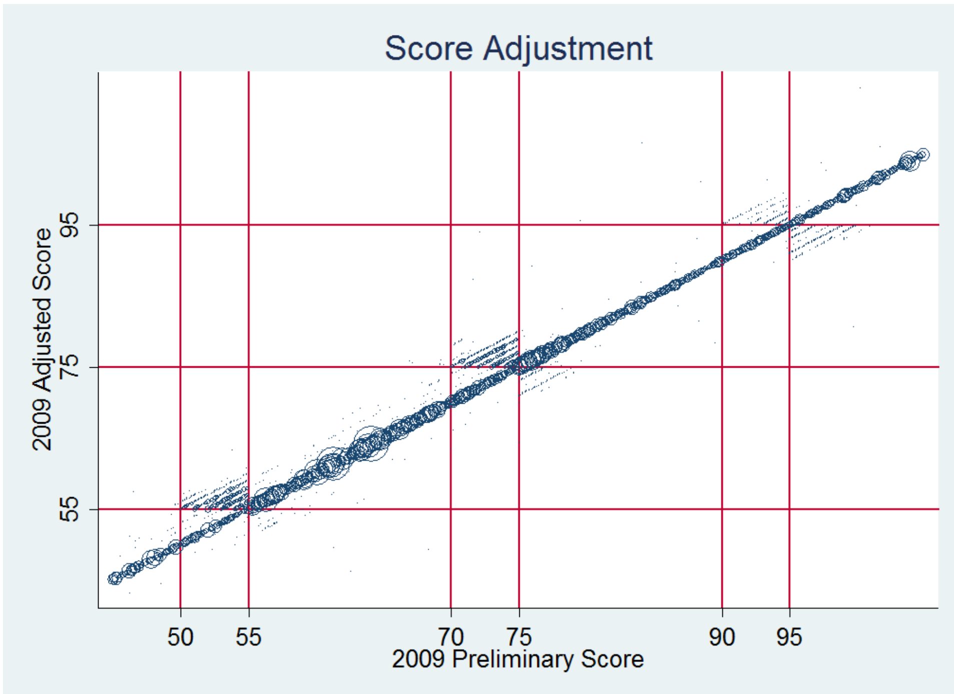
Grade 2	LTC Exp	Med Exp	Hosp Exp	Child Caregiver	Live Independently	$ Med Exp Saved / $ LTC Spent
Committee	391.4**	-94.00	-49.03	-0.00143	-0.00664	0.2
	(49.91)	(86.28)	(85.78)	(0.00735)	(0.00810)	
Treatment	1,400**	-432.8**	-419.9**	-0.0321**	-0.00845	0.3
	(45.24)	(79.39)	(79.33)	(0.00655)	(0.00708)	
Individuals	49,930	49,930	49,930	49,923	48,194	
R-squared	0.623	0.012	0.010	0.029	0.057	

Post-period regression coefficients from differences-in-differences estimation with individual fixed effects, year fixed effects, and bin-specific linear time trends. Committee = Preliminary Score in [70,75), Treatment = Preliminary Score in [75,80), Omitted = Preliminary Score in [65,70). Robust standard errors in parentheses.
** p<0.01, * p<0.05, + p<0.1

Grade 1	LTC Exp	Med Exp	Hosp Exp	Child Caregiver	Live Independently	$ Med Exp Saved / $ LTC Spent
Committee	-93.11	-326.6*	-307.6*	0.0141	-0.00297	-3.5
	(93.09)	(153.7)	(156.7)	(0.0123)	(0.0125)	
Treatment	303.4**	-589.8**	-546.4**	0.0348**	0.00128	2.0
	(95.48)	(159.0)	(161.6)	(0.0121)	(0.0130)	
Individuals	17,490	17,490	17,490	17,481	16,742	
R-squared	0.631	0.013	0.012	0.049	0.064	

Post-period regression coefficients from differences-in-differences estimation with individual fixed effects, year fixed effects, and bin-specific linear time trends. Committee = Preliminary Score in [90,95), Treatment = Preliminary Score in [95,100), Omitted = Preliminary Score in [85,90). Robust standard errors in parentheses.
** p<0.01, * p<0.05, + p<0.1

Figure 1.1: Adjusted Scores vs. Preliminary Scores, 2009

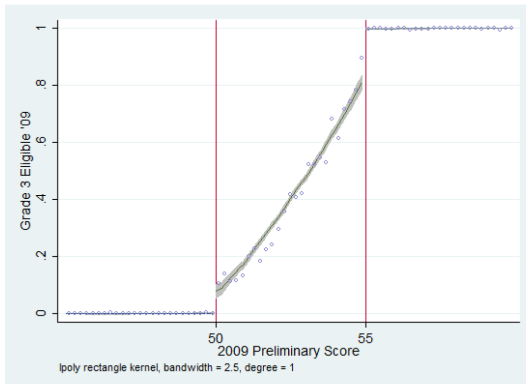


Notes: This figure plots the 2009 adjusted score against the 2009 preliminary score, for individuals whose preliminary scores fall between 45 and 105. Circle sizes correspond to the number of individuals with the associated adjusted/preliminary score combination.

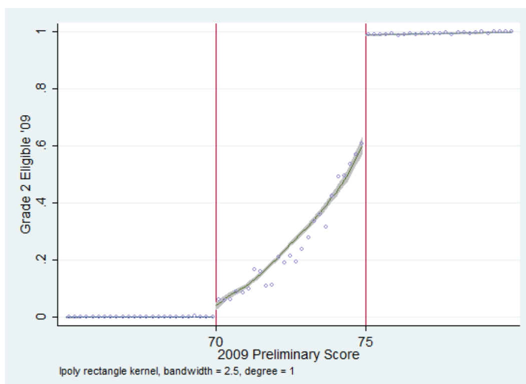
2009 adjusted score = 2009 preliminary score + committee points, where committee points $\in [-5, 5]$.

Figure 1.2: Probability of Eligibility vs. 2009 Preliminary Score

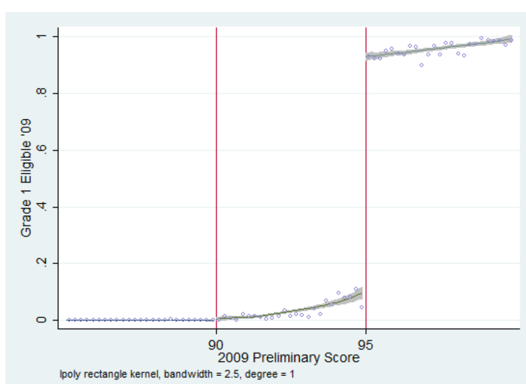
(a) Grade 3 Eligibility



(b) Grade 2 Eligibility



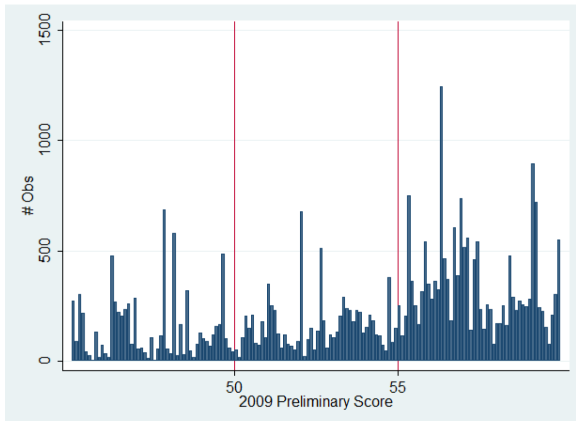
(c) Grade 1 Eligibility



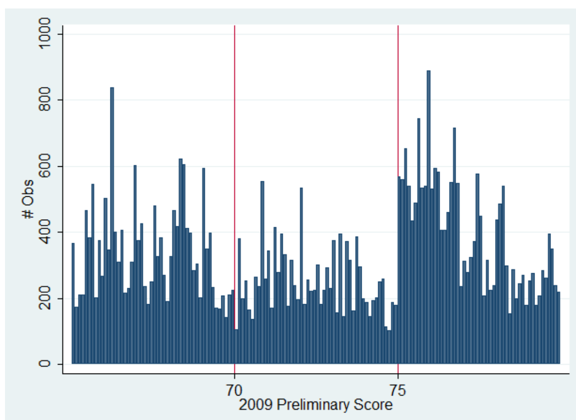
Notes: The running variable is the 2009 preliminary score. The open circles plot the mean of the dependent variable within 0.2 point bins. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 2.5 points. The shaded regions are 95 percent confidence intervals.

Figure 1.3: Histograms of Scores

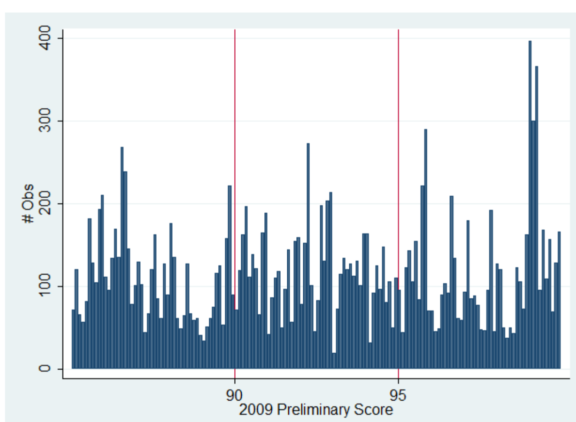
(a) Around Grade 3 Thresholds



(b) Around Grade 2 Thresholds



(c) Around Grade 1 Thresholds



Notes: 2009 preliminary score in 0.1 point bins.

Figure 1.4: Impact of $\uparrow s_H$ on A, C, and H

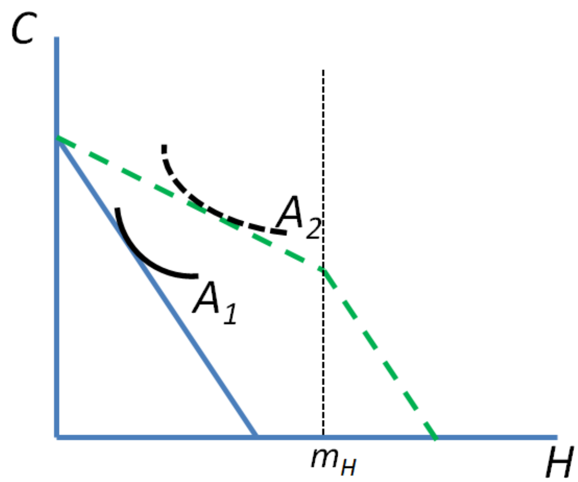


Figure 1.5: Impact of $\uparrow s_F$ on A, H, and F

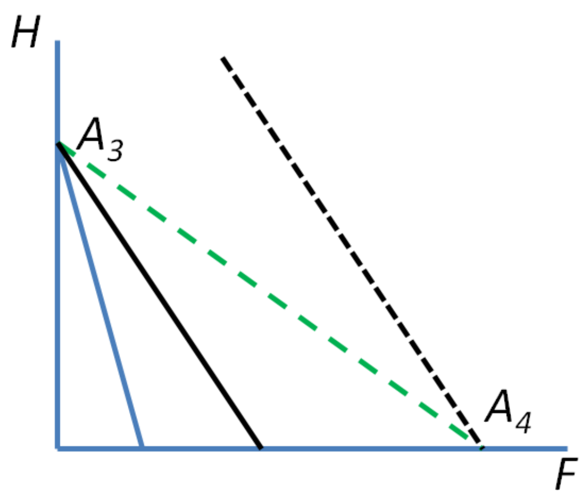


Figure 1.6: Impact of $\uparrow m_H$ on A, C, and H

(a) $H_1 < m_H$ (b) $H_1 = m_H$

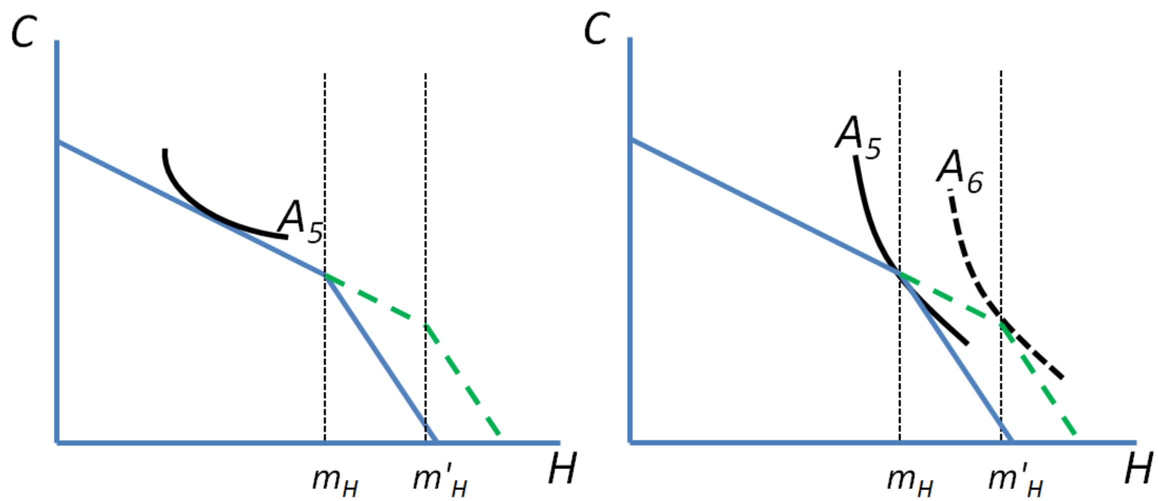
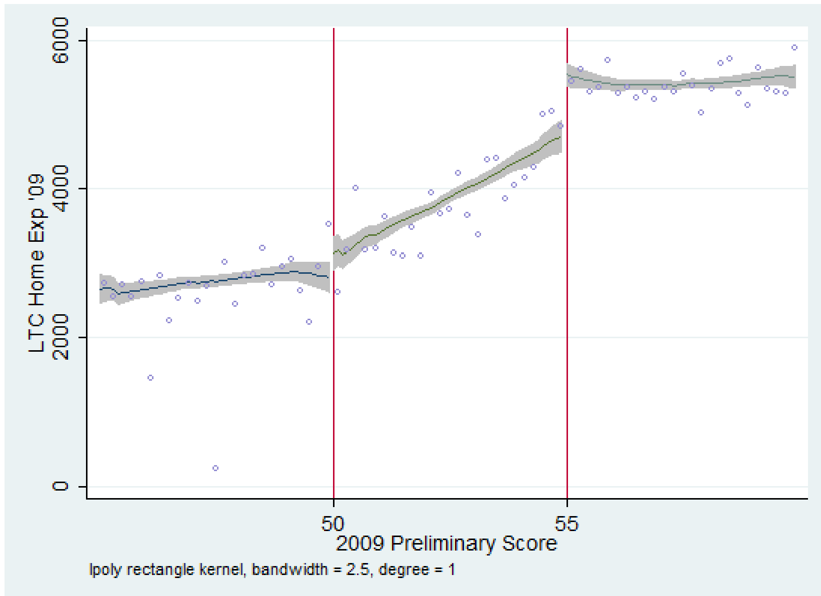
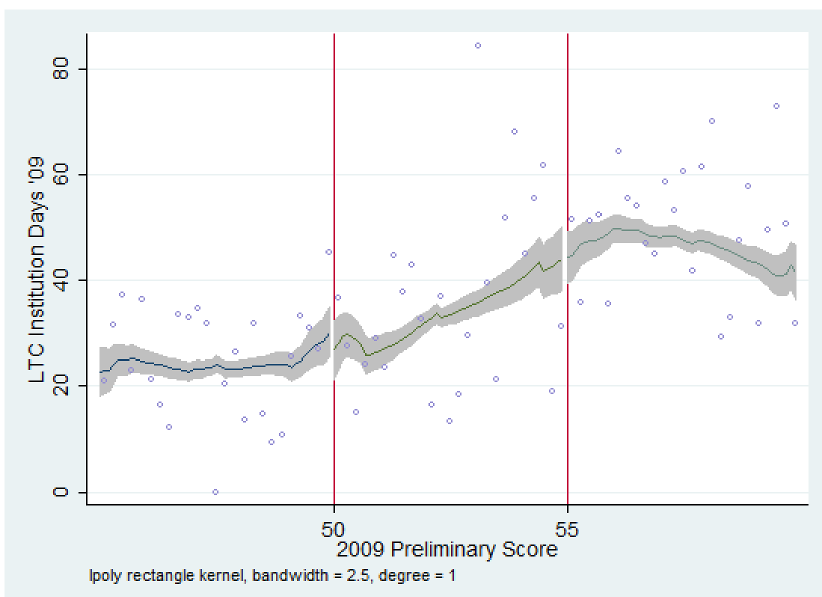


Figure 1.7: Reimbursed Formal Care Utilization vs. Preliminary Score Around Grade 3

(a) Home Care Expenditures



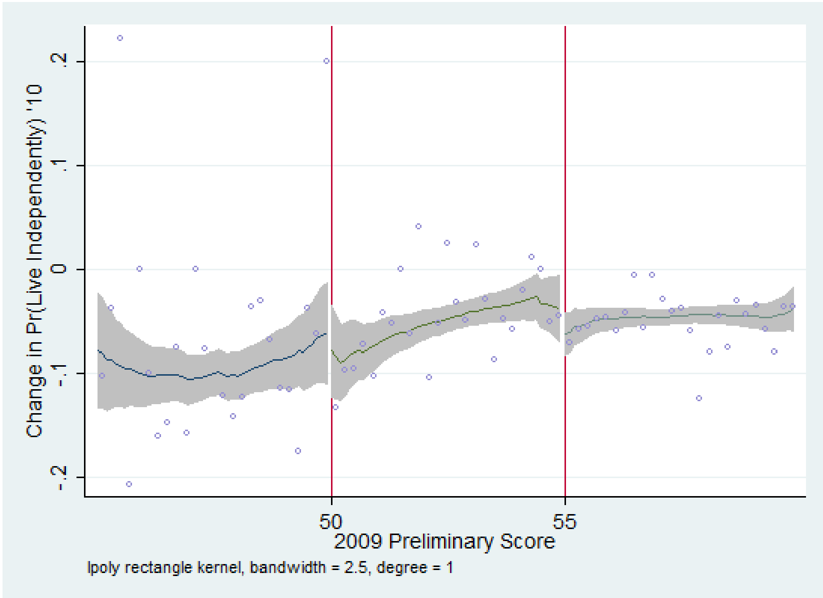
(b) Institutional Days



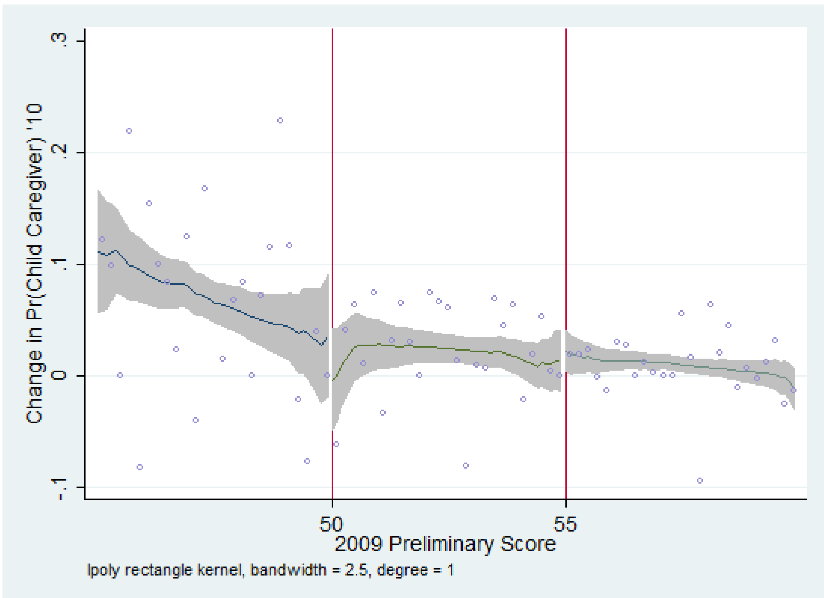
Notes: The running variable is the 2009 preliminary score. The open circles plot the mean of the dependent variable within 0.2 point bins. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 2.5 points. The shaded regions are 95 percent confidence intervals.

Figure 1.8: Change in Informal Care vs. Preliminary Score Around Grade 3

(a) Δ Pr(Live Independently)



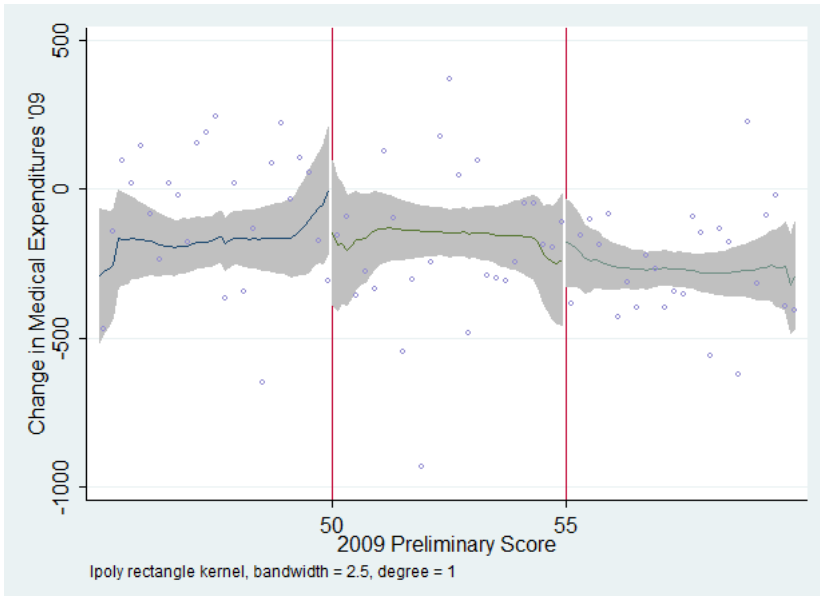
(b) Δ Pr(Child Caregiver)



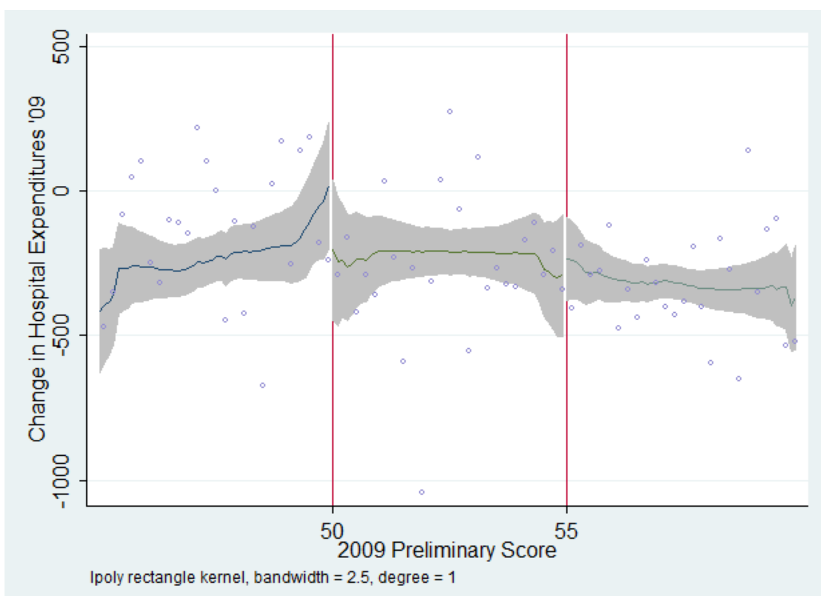
Notes: The running variable is the 2009 preliminary score. The open circles plot the mean of the dependent variable within 0.2 point bins. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 2.5 points. The shaded regions are 95 percent confidence intervals.

Figure 1.9: Change in Medical Utilization vs. Preliminary Score Around Grade 3

(a) Δ Medical Expenses



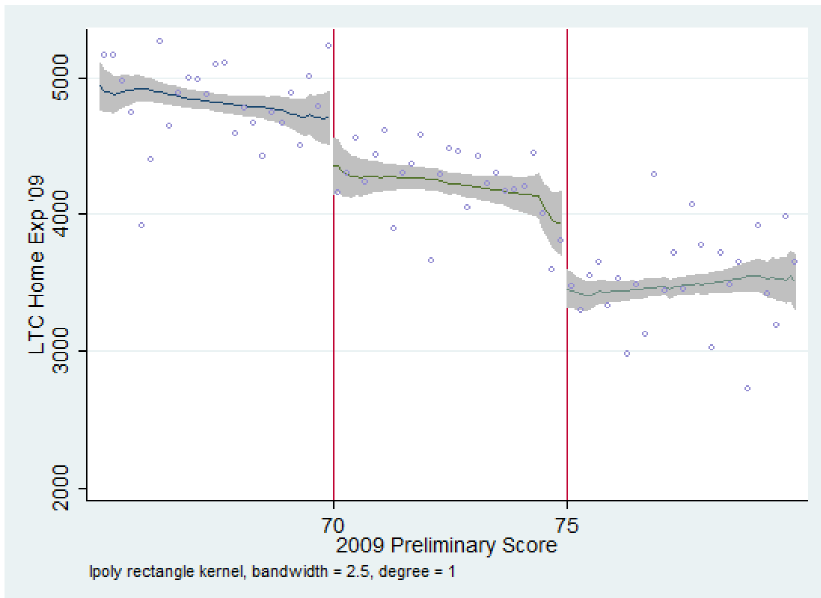
(b) Δ Hospital Expenses



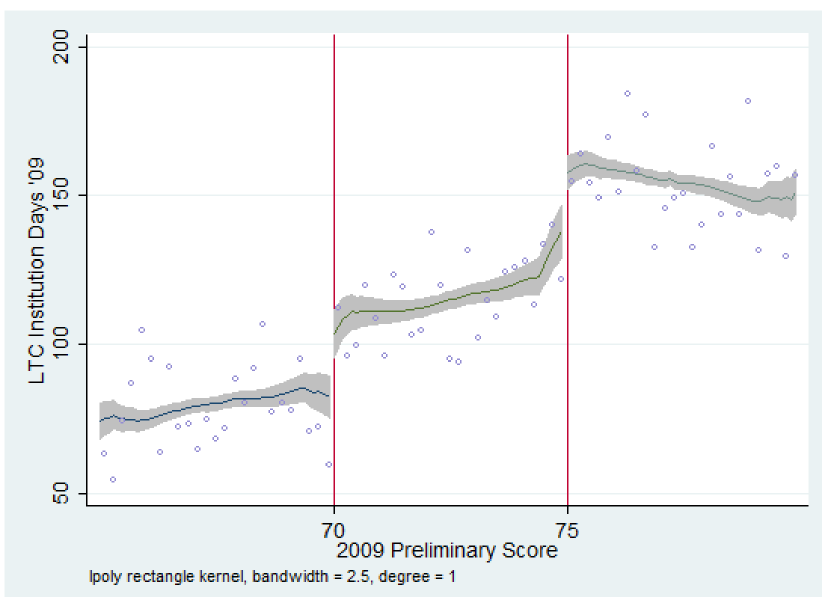
Notes: The running variable is the 2009 preliminary score. The open circles plot the mean of the dependent variable within 0.2 point bins. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 2.5 points. The shaded regions are 95 percent confidence intervals.

Figure 1.10: Reimbursed Formal Care Utilization vs. Preliminary Score Around Grade 2

(a) Home Care Expenditures



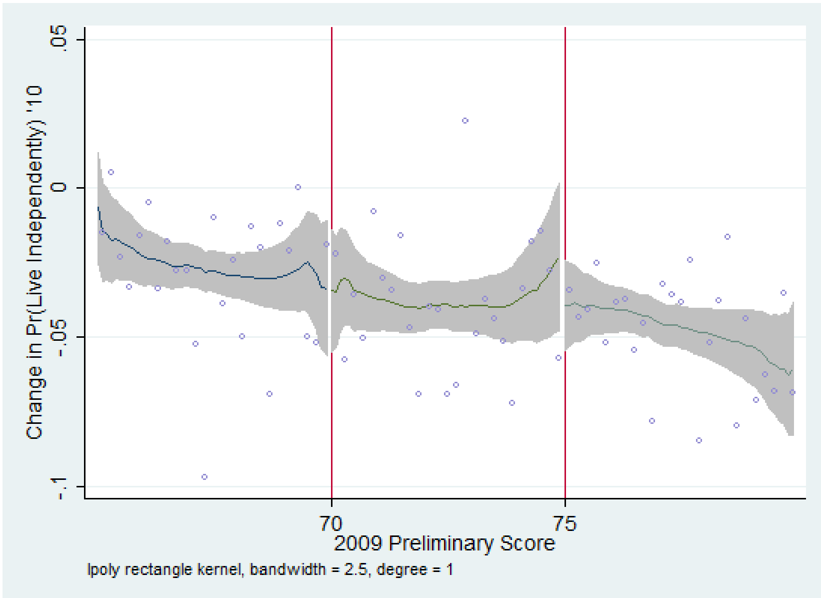
(b) Institutional Days



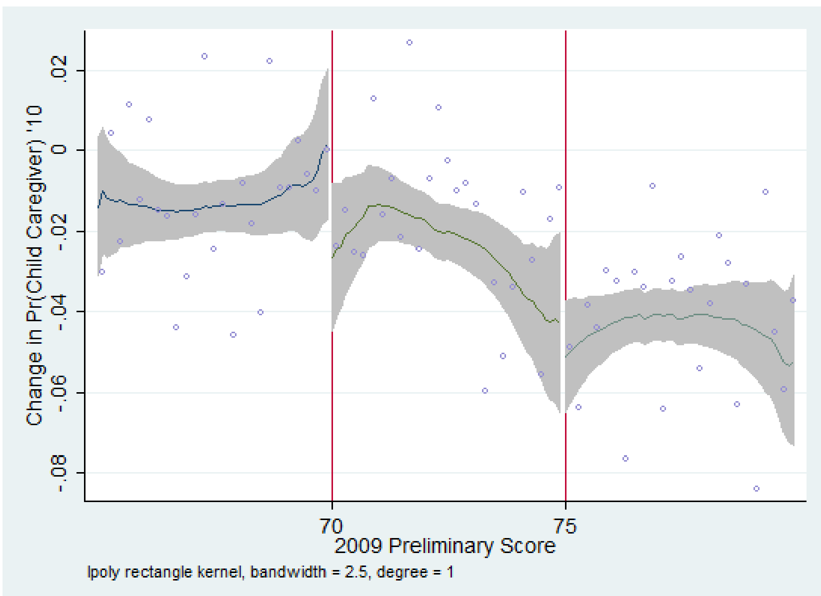
Notes: The running variable is the 2009 preliminary score. The open circles plot the mean of the dependent variable within 0.2 point bins. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 2.5 points. The shaded regions are 95 percent confidence intervals.

Figure 1.11: Change in Informal Care vs. Preliminary Score Around Grade 2

(a) Δ Pr(Live Independently)



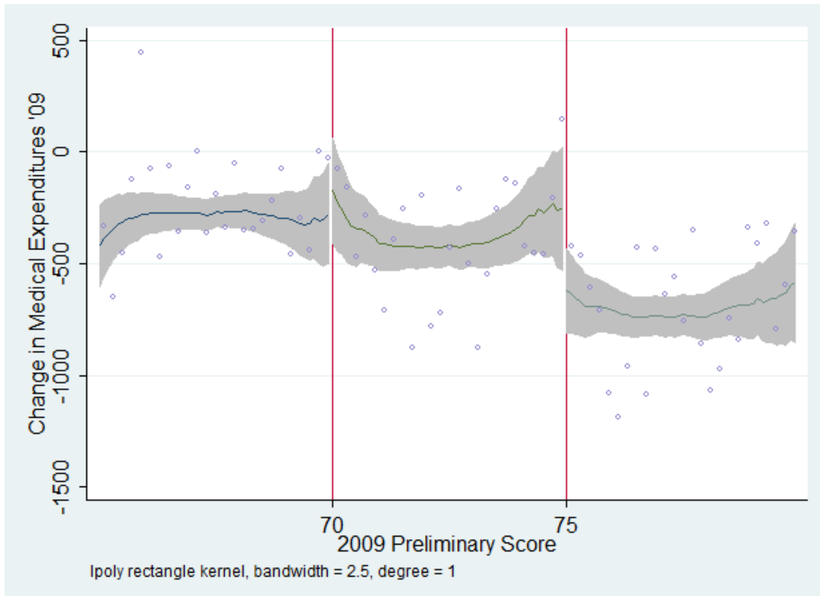
(b) Δ Pr(Child Caregiver)



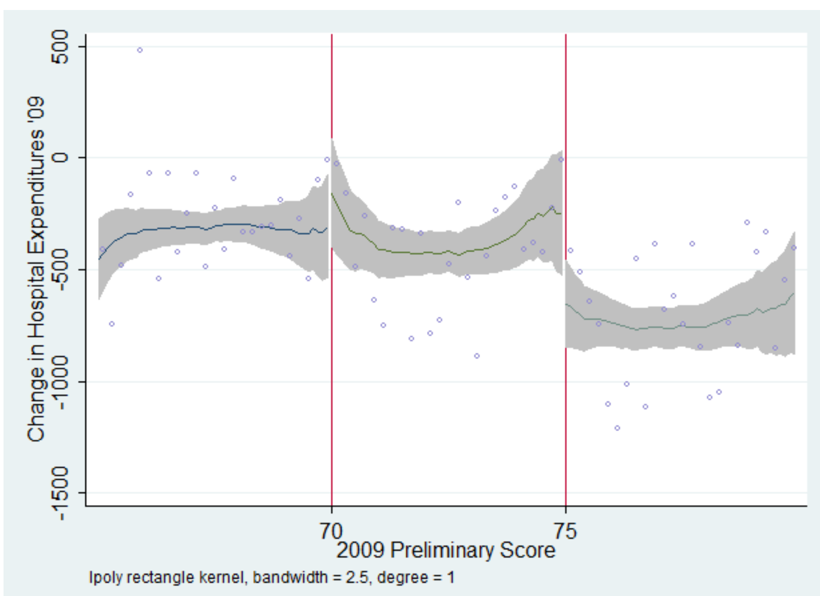
Notes: The running variable is the 2009 preliminary score. The open circles plot the mean of the dependent variable within 0.2 point bins. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 2.5 points. The shaded regions are 95 percent confidence intervals.

Figure 1.12: Change in Medical Utilization vs. Preliminary Score Around Grade 2

(a) Δ Medical Expenses



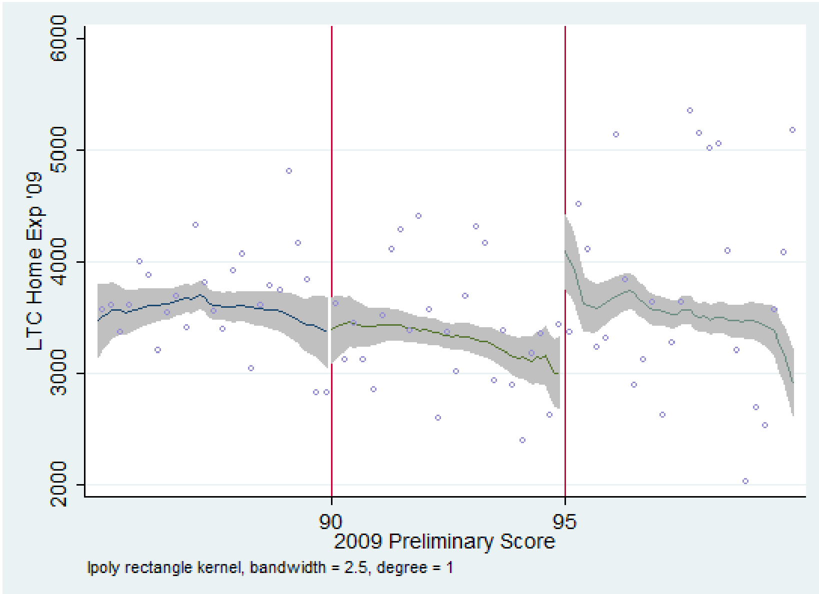
(b) Δ Hospital Expenses



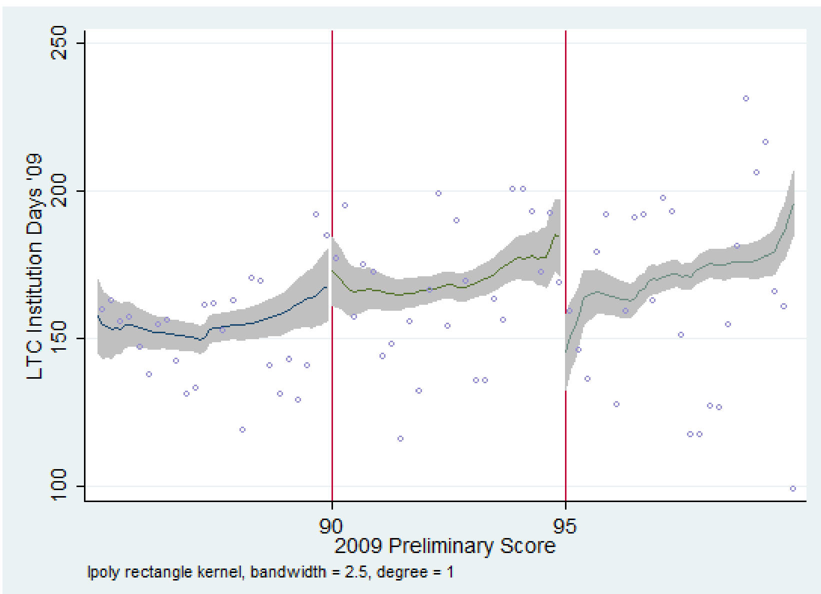
Notes: The running variable is the 2009 preliminary score. The open circles plot the mean of the dependent variable within 0.2 point bins. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 2.5 points. The shaded regions are 95 percent confidence intervals.

Figure 1.13: Reimbursed Formal Care Utilization vs. Preliminary Score Around Grade 1

(a) Home Care Expenditures



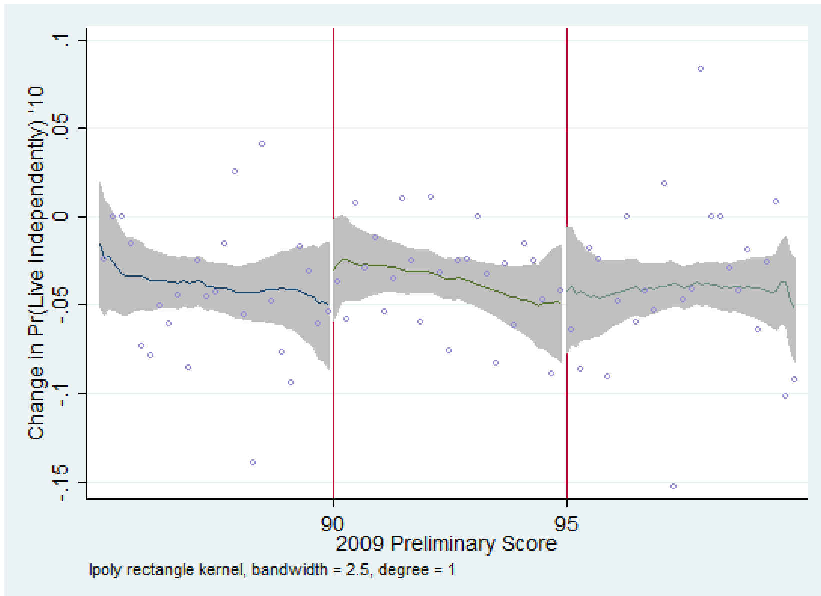
(b) Institutional Days



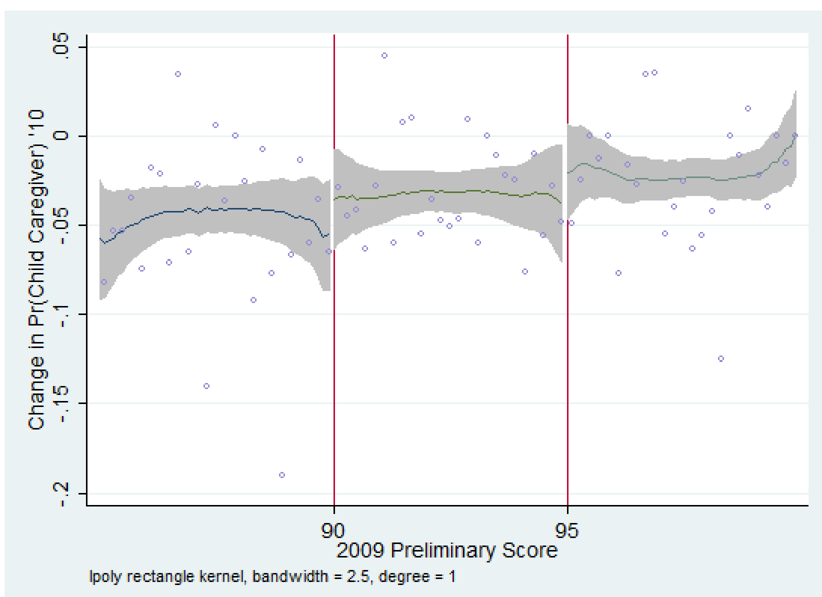
Notes: The running variable is the 2009 preliminary score. The open circles plot the mean of the dependent variable within 0.2 point bins. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 2.5 points. The shaded regions are 95 percent confidence intervals.

Figure 1.14: Change in Informal Care vs. Preliminary Score Around Grade 1

(a) Δ Pr(Live Independently)



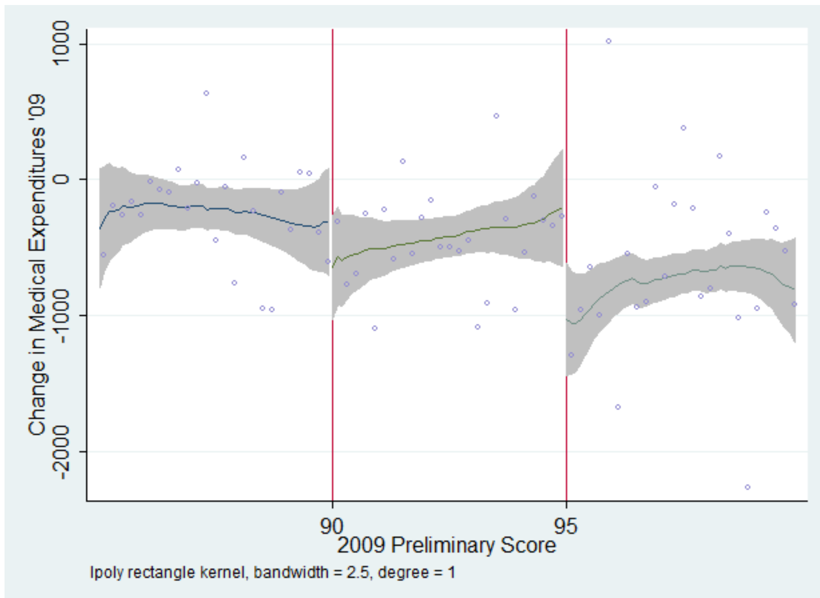
(b) Δ Pr(Child Caregiver)



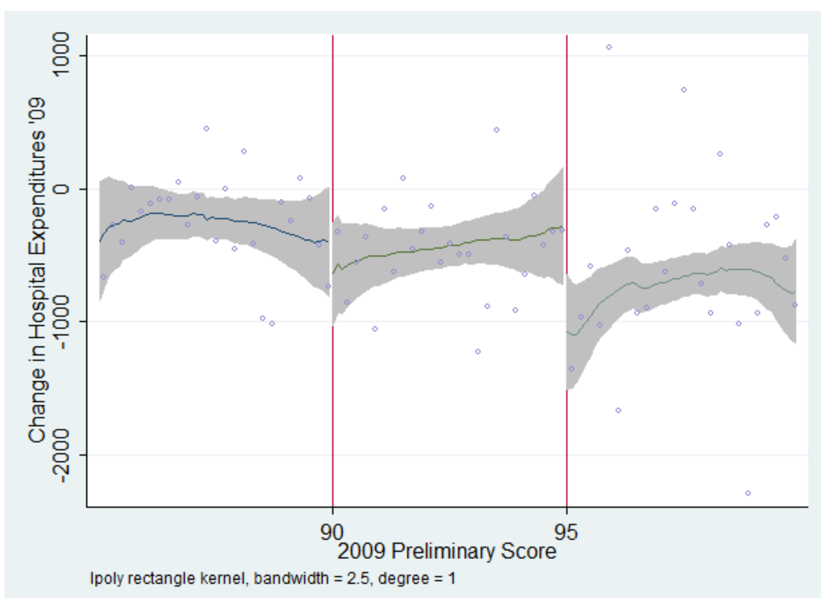
Notes: The running variable is the 2009 preliminary score. The open circles plot the mean of the dependent variable within 0.2 point bins. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 2.5 points. The shaded regions are 95 percent confidence intervals.

Figure 1.15: Change in Medical Utilization vs. Preliminary Score Around Grade 1

(a) Δ Medical Expenses



(b) Δ Hospital Expenses



Notes: The running variable is the 2009 preliminary score. The open circles plot the mean of the dependent variable within 0.2 point bins. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 2.5 points. The shaded regions are 95 percent confidence intervals.

Chapter 2

Diabetes Insurance Mandates, Behavior, and Employment

2.1 Introduction

According to the Centers for Disease Control, diabetes is the largest and fastest growing chronic disease in the U.S. Diabetes is characterized by high blood sugar because the body is unable to make (Type I) or use (Type II) insulin, resulting in complications such as heart disease, kidney disease, amputations, and blindness. Type II, which represents the vast majority of cases is largely preventable and controllable in the early stages by adhering to a healthy diet and lifestyle.

In the late 1990s, most states passed laws requiring health insurance plans to cover medical treatments and education for diabetes. This included coverage for equipment, supplies, self-management training and education, and medications related to diabetes. Figure 2.1 contains the Kentucky statute, which is very similar to the text of other states. Table 2.1 and Figure 2.2 illustrate the years these mandates were effective in each state. These dates were determined from the websites of state legislatures and LexisNexis.[1] In cases where the effective date took place after June 30, it was coded as effective the following year.

There are two dimensions where we expect these mandates to have a significant impact and which we analyze. First, we assess the effect of these mandates on individual behavior. To the extent that medical treatments and healthy behavior are substitutes in preventing diabetes complications, a reduction in the price of the former due to the mandates will lead to a decrease in the latter due to the substitution effect. However, these mandates also cover education, so the theoretical impact is ambiguous and is an empirical question. Second, we assess the effect of these mandates on employment outcomes. On one hand, these mandates increase the cost of employing individuals with diabetes. On the other hand, these benefits may increase the demand for insurance and, correspondingly, their supply of labor among diabetic individuals.

[1] For the most part, these laws were effective the year following passage.

2.2 Literature Review and Contributions

This paper fits into a larger literature looking at mandated insurance benefits. These include maternity benefits (Gruber (1994)), mental health parity (see discussion in Busch and Barry (2008)), and mammography (Bitler and Carpenter (2011)). To our knowledge, there are two papers which have looked at the impact of mandated diabetes benefits, specifically. Klick and Stratmann (2007) study the impact of these mandates on BMI and find that diabetics exhibit higher BMI after the adoption of these mandates, suggesting these mandates led to moral hazard. One way in which this paper expands on this research is by assessing the pathways through which these mandates led to increases in BMI. In particular, we seek to directly measure moral hazard by assessing behavioral responses. This is important, given that BMI is a stock and so a more relevant metric immediately following the passage of these laws are the flows into and out of it. In addition to evaluating behavioral outcomes, we also refine the estimation strategy in order to achieve cleaner identification, which we discuss in Section 2.3. Lastly, we account for the timing of the effects by looking at the time these mandates were in effect, as opposed to when they were passed.[2] This is particularly relevant for BMI given that we might expect abrupt changes in behavior but these might not translate into large level changes in BMI.

Although the economic literature on diabetes is slim, findings by Kahn (1999) suggest that diagnosed diabetics have been making better choices over time, suggesting that more medicated diabetics do not engage in less healthy choices. This stands in contrast to the findings by Klick and Stratmann (2007). In this sense, our paper stands to attempt to resolve these potentially conflicting findings.

Another paper which evaluates the impact of these mandates is Li et al. (2010). They do look at flow measures by looking at the impact on diabetes preventive care. A major

[2] Klick and Stratmann (2007) focus on the years these laws were passed, versus in effect.

limitation of their paper, however, is that they do not control for two-way interactions such as state by year effects or even state time trends. Regarding outcomes, we view this work as complementary.

In addition to looking at behavioral responses, this paper contributes to the literature by assessing the impact of these mandates on employment outcomes. As has been identified by the literature, mandated benefits laws can increase the costs of employing individuals who benefit from them (Gruber (1994)). In light of this, we develop a simple conceptual framework that generates testable predictions in order to assess the impacts of these mandates.

Lastly, in looking at moral hazard in response to these mandated benefits, this paper fits more generally into the literature on offsetting behavior (see for example Peltzman (1975)).

2.3 Data

The main data used in this analysis come from the Behavioral Risk Factor Surveillance System (BRFSS), which is an annual cross-sectional survey conducted by state health departments and compiled by the Centers for Disease Control. This dataset contains survey questions on diabetes status, exercise, diet, employment, household income, and insurance. Our analysis uses data from 1994–2003. Prior years did not account for gestational diabetes and the income variable was also defined differently. Years after 2003 separated out pre-diabetes. To reduce measurement error from diagnosis of gestational diabetes among women, we limit our analysis to men. Because most individuals age 65 and older are covered by Medicare, we also restrict our sample to men under age 65 because the diabetes mandates apply to private insurance. Summary statistics are presented in Table 2.2. Our measure of exercise is a binary measure of whether the individual participated in any physical activities or exercises in the past month. Our measures of diet come from the number of servings of fruits, salad, and vegetables consumed. We convert this to a daily value and transform it by taking the log of one plus this number. Our proxy for wages is household income, which is measured in categories and converted to the mean of each category. We also use data from the Current Population Survey (CPS) when we analyze employment outcomes, as this dataset gives us better measures of employment and wages than BRFSS.

2.4 Individual Behavior

2.4.1 Conceptual Framework

The risk of diabetes related health complications can be managed by medications and healthy behavior. If we consider medications and healthy behavior to be substitutes and think of the diabetes mandates as reducing the price of medications, we would expect a reduction in healthy behavior. On the other hand, increased education as a result of the mandates may lead individuals to engage in healthier behavior. Thus, the net effect is uncertain. Moreover, previous evidence is mixed. Klick and Stratmann (2007) find that the diabetes mandates led to increases in BMI among diabetics. However, Kahn (1999) provides evidence against moral hazard among diabetics in the face of improved medical technology. We assess the net impact empirically by focusing directly on behavior.

2.4.2 Empirical Framework

We are interested in identifying the causal effects of state laws which required health insurance plans to cover medical treatments and education for diabetes. One concern is that unobserved characteristics may contribute to changes in health behavior and adoption of state mandates. For example, changes in the health care system within states may have coincided with the timing of the mandates. To estimate the effect of the state diabetes mandates on our outcomes of interest, we estimate triple differences models that identify the effects of the mandates using variation in which states adopted, when they adopted, and the populations affected. Specifically, we estimate the following:

$$\text{outcome}_{ist} = \beta \cdot [Diab_i \cdot Mand_{st}] + \eta_s \cdot \phi_t + \eta_s \cdot Diab_i + \phi_t \cdot Diab_i + \gamma \cdot Diab_i + \epsilon_{ist}, \quad (2.1)$$

where $Diab_i$ is an indicator for being diabetic, $Mand_{st}$ is an indicator for living in a mandate state in a year that the mandate is in effect, η_s is a state fixed effect, and ϕ_t is a year fixed effect. The two-way interactions allow us to control for unobserved changes across states and years that impact diabetics and non-diabetics similarly, time invariant differences between diabetics and non-diabetics within a state, and national level factors that lead to differences between diabetics and non-diabetics such as new medical advances for diabetes over time. Thus, the underlying assumption is that there are no unobserved factors that affect diabetics and non-diabetics differently within a state and year.

This triple differences analysis would not be valid if these laws had impacts on the prevalence of diabetes, particularly if the composition of those with diabetes changed as a result. To assess the extent to which this may have happened we conduct an event study analysis, looking at the prevalence of diabetes before and after states adopted the diabetes mandates. Specifically, we estimate

$$Diab_{ist} = \sum_{k=-3}^{3} \gamma_k \cdot k + \eta_s + \phi_t + \eta_s \cdot t + \epsilon_{ist},$$

where k is the number of years the mandate has been in effect (inclusive), and $\eta_s \cdot t$ captures state specific linear time trends. Figure 2.3 displays estimates of k, showing changes in diabetes prevalence relative to the year prior the law became effective. There is no evidence for an impact on diabetes prevalence. Another threat to our analysis is if these laws affected a diabetic individual's insurance status. As a result, we look at this as an outcome in our analysis in the next section.

2.4.3 Results

We present estimates of Equation 2.1 for our outcomes of interest in Table 2.3. For reference, we also include a less stringent specification that allows for the two-way diabetes by mandate-state and mandate-state by post interactions as well as state fixed effects, year fixed effects, and state linear time trends. It appears that the less stringent specification suggests that these mandates are associated with a negative impact on exercise and fruit but no impact on the consumption of salad or vegetables. However, these results are not robust to including state-year effects. Ultimately we find no statistically significant impact on exercise, diet, or BMI.

As discussed in the previous section, our estimates may be affected if these mandates affected a diabetic individual's insurance status. Table 2.3 shows that there is no impact on insurance status. Another concern is that our results are sensitive to specifying the timing of the effect of the mandates. To assess the sensitivity of our results to this assumption we reestimate Equation 2.1 where $Mand_{st}$ equals 1 in years one year prior to the effective date and after (one year lead), as well as when $Mand_{st}$ equals 0 until the year following the effective date (one year lag). Table 2.4 indicates that our findings are not sensitive to how we specify the timing of the mandates.

2.5 Employment Outcomes

2.5.1 Conceptual Framework

We can think of these mandates as increasing the costs of insuring a particular group—diabetics and individuals at risk of diabetes. To the extent that these costs are passed on to the employer, in a competitive market wages would fall to offset the cost of the benefit. While diabetes is not easily identifiable, correlates of diabetes, such as BMI and race, are.[3] Thus, we might expect wages of those with these characteristics to fall as a result of the passage of these mandates. One issue is that there is less scope for relative adjustment of wages because of anti-discrimination rules. As a result, we might expect an effect on the employment of these individuals. In this respect, we follow an approach similar to Gruber (1994) who studies the labor market effects of mandated maternity benefits, which raise the costs of employing a demographically identifiable group (women of childbearing age). In this case, we assess changes in wages and employment among groups with higher prevalence of diabetes, specifically those with high BMI and minorities.

In addition to the effects on labor demand, however, we can also think of these mandates as increasing the labor supply of individuals with diabetes. The fact that these two groups—individuals who are more likely to have diabetes based on observable characteristics, and individuals who do have diabetes (which we assume is unobservable to the employer)—do not perfectly coincide leads to testable predictions, summarized in Table 2.5. Specifically, because minorities are more likely to have diabetes, these mandates increase the expected cost of employing a minority. At the margin, this will decrease labor demand for minorities relative to white individuals. However, not all minorities are affected equally by these mandates. Minorities with diabetes will derive some benefit from the additional coverage, but those without diabetes will not, leading labor supply to increase for for the former but not the latter. The net impact of labor demand and supply for minority diabetics will then lead to a decrease in the wages of these individuals, with an ambiguous impact on employment. For minority non-diabetics, the effect is due solely to decreased labor demand, which leads to

[3]Minorities have higher prevalence of diabetes than whites. In our data, among men under age 65, blacks have a prevalence rate of 7.5% while whites have a prevalence rate of 4.1%. Diabetes is also more prevalent among overweight individuals. In our data, diabetes prevalence among obese individuals is 5.5% versus 2.1% for those who are not.

a decrease in both wages and employment. For white diabetics, the impact is primarily from increased labor supply which results in a decrease in the wage but an increase in employment. These differential impacts for different groups of individuals motivate our empirical framework.

2.5.2 Empirical Framework

The empirical framework for looking at employment outcomes is motivated by the fact that individuals who are more likely to have diabetes based on observable characteristics are not the same as individuals who actually have diabetes. As in Table 2.5, changes in labor supply and labor demand depend on both race and diabetes status, which lead to different predictions for wage and employment along these different dimensions. This motivates a triple difference analysis where we compare individuals of different race and diabetes status, in mandate and non-mandate states, before and after the diabetes mandates take effect. The model we estimate is:

$$\text{outcome}_{ist} = \beta \cdot [Group_i \cdot Mand_{st}] + \eta_s \cdot \phi_t + \eta_s \cdot Group_i + \phi_t \cdot Group_i + \gamma \cdot Group_i + \epsilon_{ist}, \quad (2.2)$$

where $Group_i$ is “black and diabetic”, “black and non-diabetic”, or “white and diabetic”. In all cases, the omitted group is “white and non-diabetic”.

2.5.3 Results from BRFSS Data

We present estimates of β in Equation 2.2 in Table 2.6. As before, we also present a less stringent specification that allows for the two-way group by mandate-state and mandate-state by post interactions as well as state fixed effects, year fixed effects, and state linear time trends. There appear to be no statistically significant effects on the income and employment of minority and/or diabetic individuals relative to white non-diabetics. The results for employment are particularly precise, ruling out impacts of more than 1% among black non-diabetics and white diabetics. We also check the sensitivity of our results to the timing of the mandates. Table 2.7 indicates our findings are not sensitive to how we specify the timing of the mandate.

2.5.4 Results from CPS Data

One limitation of the BRFSS data is that it does not contain measures of wages. To overcome this issue as well as verify our results in another dataset we turn to the CPS. One limitation of the CPS is that it does not contain information on diabetes status. It does contain information on race, however. Moreover, our conceptual framework generates an unambiguous prediction for (decreased) wages among blacks relative to whites, regardless of diabetes status. Thus, we estimate a triple differences model similar to Equation 2.1, except that $Diab_i$ is replaced with $Black_i$. Estimates are presented in Table 2.8. Even though our conceptual framework predicts an unambiguous decrease in wages for blacks relative to whites, we find no statistically significant effect. There is a surprising positive effect on wages, but this is not statistically significant at the 5% level.

2.6 Discussion and Conclusion

In summary, we find that diabetes insurance mandates had no impact on healthy behavior, as would be expected if a reduction in the cost of managing diabetes led to moral hazard. Moreover, despite the increased costs associated with employing diabetics, we find no impact on the employment outcomes of diabetics and groups with higher prevalence of diabetes.

One potential explanation for these results is that these state mandates do not affect all insurance plans. Specifically, the Employee Retirement Income Security Act of 1974 (ERISA) exempts employers who self-insure from these laws. As discussed in Bitler and

Carpenter (2011), around 30% of workers were enrolled in non-self-insured plans in 2000. In light of this, we are able to rule out effects larger than 9.8% (=(-0.00305-1.96*0.0166)/0.3) for exercising within the past month. Our estimates on unemployment of blacks relative to whites are even more precise, ruling out effects larger than 2% (=(-0.00127-1.96*0.00371)).

Table 2.1: Diabetes Mandates—Year in Effect

1994	1996	1997	1998	1999	2000	2001	2002	Never
New York	Florida	Maine	Arkansas	Arizona	Alaska	Delaware	Hawaii	Alabama
	New Jersey	Minnesota	Connecticut	Colorado	California	Massachusetts	Montana	Idaho
		Oklahoma	Indiana	Georgia	Iowa	Michigan	Oregon	Missouri
		Rhode Island	Louisiana	Illinois	Nebraska	Utah	Wyoming	North Dakota
		West Virginia	Maryland	Kansas	South Carolina			Ohio
			Nevada	Kentucky	South Dakota			
			New Hampshire	Mississippi				
			New Mexico	Pennsylvania				
			North Carolina	Virginia				
			Tennessee					
			Texas					
			Vermont					
			Washington					
			Wisconsin					

Table 2.2: Summary Statistics

	Total	Non-Diabetic	Diabetic
Diabetic	0.04		
BMI	26.9	26.8	29.9
Exercised in Past Month	0.77	0.78	0.67
Fruit Servings/Day	0.66	0.66	0.81
Salad Servings/Day	0.44	0.44	0.51
Vegetable Servings/Day	1.17	1.16	1.29
Insured	0.84	0.84	0.87
Employed	0.95	0.95	0.92
Income (000s)	41,862	42,063	37,563
Mandate in Effect	0.57	0.57	0.65
Age	40.7	40.2	50.8
Education Level	4.8	4.8	4.6
N	551,870	527,110	24,760

Notes: Behavioral Risk Factor Surveillance System. Males below age 65. N is maximum number of observations.

Table 2.3: Effect of Diabetes Mandates on Healthy Behavior

Controls	State FE, Year FE, State Linear Time Trends						State x Year FE, Diab x State FE, Diab x Year FE					
Dependent Variables	**BMI**	**Exercise**	**Fruit**	**Salad**	**Vegetables**	**Insured**	**BMI**	**Exercise**	**Fruit**	**Salad**	**Vegetables**	**Insured**
$Diab_i$ x $Mand_{st}$	0.558**	-0.0258*	-0.0373**	0.000563	0.00732	0.00392	0.203	-0.00305	-0.0210	-0.00115	-0.00695	-0.00503
	(0.0917)	(0.0102)	(0.00737)	(0.00552)	(0.00798)	(0.00577)	(0.129)	(0.0166)	(0.0131)	(0.00904)	(0.0117)	(0.00899)
Observations	539,447	440,067	374,770	376,269	373,914	549,912	539,447	440,067	374,770	376,269	373,914	549,912
R-squared	0.033	0.020	0.019	0.022	0.033	0.011	0.034	0.021	0.020	0.022	0.031	0.012

Notes: This table reports estimates of β from Equation 2.1. $Mand_{st}$ indicates a diabetes mandate is in effect in the individual's state. Fruit, salad, and vegetables are the log of the number of daily servings plus one. Exercise is a binary measure of whether the individual exercised in the past month. Robust standard errors clustered by state are in parantheses.

Table 2.4: Effect of Diabetes Mandates on Healthy Behavior—Sensitivity to Timing of Mandates

Definition of $Mand_{st}$	**Mandate in Effect No Later Than Next Year**						**Mandate in Effect Last Year or Earlier**					
Dependent Variables	**BMI**	**Exercise**	**Fruit**	**Salad**	**Vegetables**	**Insured**	**BMI**	**Exercise**	**Fruit**	**Salad**	**Vegetables**	**Insured**
$Diab_i$ x $Mand_{st}$	0.0554	-0.00905	0.00341	-0.00840	0.0166	-0.0121	0.151	-0.00401	-0.0117	0.00644	-0.00805	-0.00815
	(0.161)	(0.0145)	(0.0120)	(0.00897)	(0.0119)	(0.00770)	(0.149)	(0.0153)	(0.0152)	(0.00990)	(0.0118)	(0.00871)
Observations	539,447	440,067	374,770	376,269	373,914	549,912	539,447	440,067	374,770	376,269	373,914	549,912
R-squared	0.034	0.021	0.020	0.022	0.031	0.011	0.034	0.021	0.020	0.022	0.031	0.012

Notes: This table reports estimates of β from Equation 2.1. Fruit, salad, and vegetables are the log of the number of daily servings plus one. Exercise is a binary measure of whether the individual exercised in the past month. Controls include state by year, diabetic by state, and diabetic by year fixed effects. Robust standard errors clustered by state are in parentheses.

Table 2.5: Labor Supply and Demand, by Race and Diabetes Status

	Labor Supply	Labor Demand	Wage	Employment
Minority diabetics	↑	↓	↓	?
Minority non-diabetics	–	↓	↓	↓
White diabetics	↑	–	↓	↑
White non-diabetics	–	–	–	–

Table 2.6: Effect of Diabetes Mandates on Employment Outcomes

Group	Black and Diabetic		Black and Non-Diabetic		White and Diabetic	
Dep. Variable	Employed	HH Inc	Employed	HH Inc	Employed	HH Inc
$Group_i$ x $Mand_{st}$	0.0240	-775.9	0.00274	329.0	0.000281	551.7
	(0.0271)	(1,629)	(0.00735)	(655.4)	(0.00929)	(578.1)
Observations	375,564	388,078	402,992	415,548	386,090	402,129
R-squared	0.008	0.064	0.010	0.074	0.007	0.062

Notes: This table reports estimates of β from Equation 2.2. The reference group in each case is white and non-diabetic. $Mand_{st}$ indicates a diabetes mandate is in effect in the individual's state. Controls include state by year, group by state, and group by year fixed effects. Robust standard errors clustered by state are in parentheses.

Table 2.7: Effect of Diabetes Mandates on Employment Outcomes

Definition of $Mand_{st}$	Mandate in Effect No Later Than Next Year						Mandate in Effect Last Year or Earlier					
Group	Black and Diabetic		Black and Non-Diabetic		White and Diabetic		Black and Diabetic		Black and Non-Diabetic		White and Diabetic	
Dep. Variable	Employed	HH Inc	Employed	HH Inc	Employed	HH Inc	Employed	HH Inc	Employed	HH Inc	Employed	HH Inc
$Group_i$ x $Mand_{st}$	-0.0280	1,125	0.00158	-825.4	-0.0177+	59.91	-0.0131	1,340	0.00687	-266.9	0.0193+	42.42
	(0.0385)	(1,636)	(0.00588)	(539.2)	(0.00957)	(535.5)	(0.0321)	(1,603)	(0.00444)	(527.8)	(0.0107)	(612.4)
Observations	375,564	388,078	402,992	415,548	386,090	402,129	375,564	388,078	402,992	415,548	386,090	402,129
R-squared	0.008	0.064	0.010	0.074	0.007	0.062	0.008	0.064	0.010	0.074	0.007	0.062

Notes: This table reports estimates of β from Equation 2.2. The reference group in each case is white and non-diabetic. Controls include state by year, group by state, and group by year fixed effects. Robust standard errors clustered by state are in parentheses. ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

Table 2.8: Effect of Diabetes Mandates on Employment Outcomes

Dep. Variable	Unemployed	Log(Wage)
$Black_i$ x $Mand_{st}$	-0.00127	0.0439+
	(0.00371)	(0.0220)
Observations	625,608	615,322
R-squared	0.010	0.030

Notes: This table reports estimates of β from Equation 2.1, where $Diab_i$ is replaced with $Black_i$. $Mand_{st}$ indicates a diabetes mandate is in effect in the individual's state. Controls include state by year, black by state, and black by year fixed effects. Robust standard errors clustered by state are in parentheses. ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

Figure 2.1: Kentucky Diabetes Mandate

304.17A-148 Coverage for diabetes.

(1) All health benefit plans issued or renewed on or after July 15, 1998, shall provide coverage for equipment, supplies, outpatient self-management training and education, including medical nutrition therapy, and all medications necessary for the treatment of insulin-dependent diabetes, insulin-using diabetes, gestational diabetes, and noninsulin-using diabetes if prescribed by a health care provider legally authorized to prescribe the items.

(2) Diabetes outpatient self-management training and education shall be provided by a certified, registered, or licensed health care professional with expertise in diabetes, as deemed necessary by a health care provider.

(3) (a) The benefits provided in this section shall be subject to the same annual deductibles or coinsurance established for all other covered benefits within a given health benefit plan.

(b) Private third-party payors may not reduce or eliminate coverage due to the requirements of this section.

Figure 2.2: Diabetes Mandates—Year in Effect

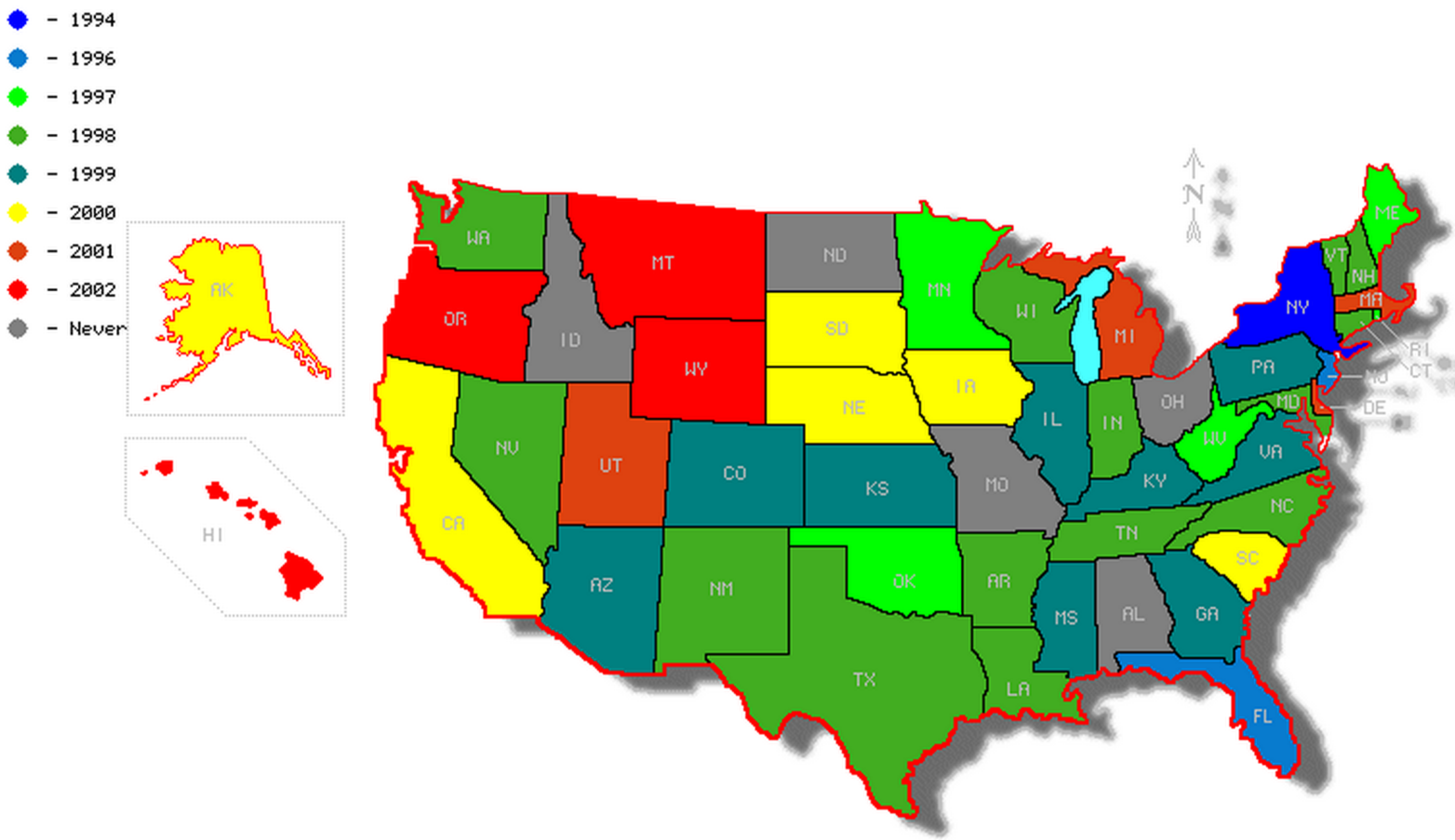
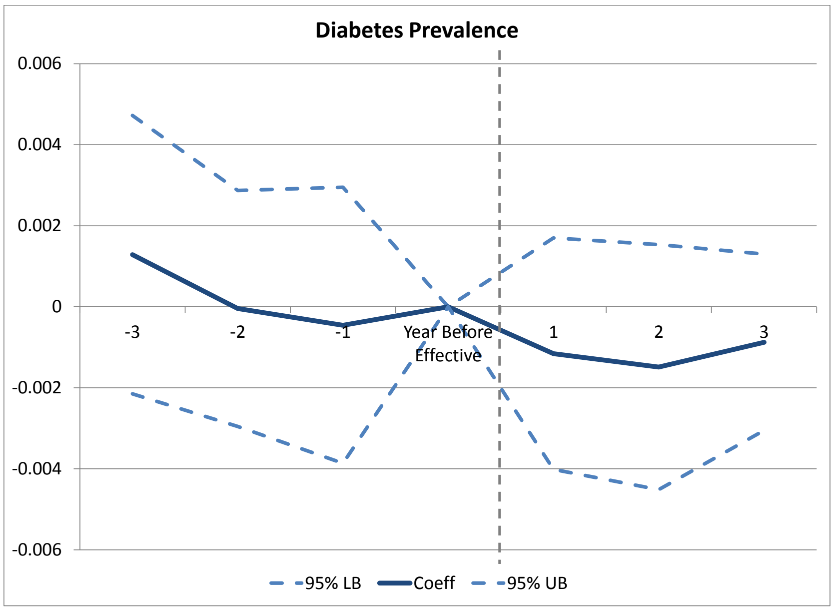


Figure 2.3: Diabetes Prevalence Before and After Mandates Took Effect



Chapter 3

Is Knowing Half the Battle? The Case of Health Screenings

with Hyuncheol Kim

3.1 Introduction

According to the Centers for Disease Control, diabetes is the largest and fastest growing chronic disease in the U.S. Diabetes is characterized by high blood sugar because the body is unable to make or use insulin, resulting in complications such as heart disease, kidney disease, amputations, and blindness. It is frequently not diagnosed until complications appear and as a result almost one-third of all people with diabetes may be undiagnosed (CDC (2011)). Moreover, there is evidence for the benefits of early treatment of diabetes or prediabetes diagnosed through usual clinical care. These facts suggest the possible benefits from screenings of asymptomatic individuals. This is the first study based on a quasi-experimental design to analyze the impact of screening for diabetes in the general population. While we do not assess the effectiveness of screening versus not screening, we do assess the impact of information obtained from screening.

The motivation behind screening in general is that early detection and intervention can increase the likelihood of reducing complications from disease or eliminating the disease altogether. However, the effectiveness of screening depends on many factors, including the reliability of the relevant diagnostic test, the effectiveness of treatments in terms of both efficacy and quality of life, and costs. An example of a disease for which the potential risks are thought to outweigh the benefits of screening is prostate cancer. This is due to false positives, benign cancers, and serious side effects (USPSTF (2008)). In the case of diabetes, the U.S. Preventive Services Task Force currently makes no recommendation for diabetes screening for individuals without high blood pressure, due to the lack of sufficient evidence.

In the case of diabetes, for which exercise and diet play a substantial role in the management of the disease and prevention of complications, an important component determining the benefit of screening is the degree to which results from the screening lead individuals to undergo changes in behavior. In light of this, whether and how individuals process the information from screenings is an important question. Much literature has been devoted to understanding the role of education in the production of health (Cutler and Lleras-Muney (2006), Grossman (2006), Lange (2011), Lleras-Muney (2005)). This paper is complementary to this literature by accounting for factors such as education that are correlated with both the attainment of information and their responses to that information. In addition to assessing patients' responses and outcomes due to that information we also study whether such responses vary according to a proxy for education.

We address these questions by exploiting a unique program where individuals undergo screening for various health measures, including blood sugar for diabetes, and then receive notification of their health status classification—either "normal", "risk group", or "suspected disease". These classifications vary discontinuously at different blood sugar thresholds which enable us to assess individuals' responses to those classifications while controlling for unobservable factors correlated with both the attainment of information and responses to that information. We assess longer term outcomes which are expected to be affected by screening—mortality, medical expenditures, and hospital days. We also assess intermediate outcomes, including follow-ups and future screening use, in order to shed light on possible mechanisms. Lastly, we assess whether responsiveness varies by education, in order to shed light on the role of education in processing information.

We find that individuals who are classified as "risk group" have no different mortality and medical outcomes within five years of screening than those who are classified as "normal". Similarly, we find that individuals who are classified as "disease suspected" have no different mortality and medical outcomes within five years of screening than those who are classified as "risk group". In attempting to shed light on mechanisms, we find no statistically significant differences in physician visits or screening follows-ups among the same pairwise comparison groups. However, we do find substantial undertaking of secondary examinations among those in the "risk group", who were encouraged and eligible to do so. When assessing differences in this outcome by insurance contribution (a proxy for education), we see that those in the highest quintile respond less than those in the other quintiles, which may be explained by the fact that more educated (higher insurance contribution) individuals respond less to the classification because they have already largely incorporated this information from the blood sugar measure itself.

The remainder of this paper is organized as follows. Section 3.2 describes the institutional context and the screening program which creates the setting for this analysis. Section 3.3 describes the data and Section 3.4 the empirical framework. Section 3.5 presents the results. Section 3.6 concludes and provides direction for future work.

3.2 Institutional Details

Korea provides universal health care. Individuals are covered either by National Health Insurance (NHI) or Medical Care Assistance (MCA), though both programs are overseen by the National Health Insurance Corporation (NHIC). The primary distinction between NHI and MCA is that the latter serves poor individuals.

NHI operates the National Health Screening Program (NHSP), which since 1995 has provided general screening services to people ages 40 and over free of charge every two years. People born in odd-numbered years are encouraged to undergo screening in odd-numbered years and vice versa. NHSP consists of the recording of medical history; measuring of height, weight, blood pressure, vision, and hearing; chest X-ray; urine sample; blood test, including hemoglobin, cholesterol, and rGTP; oral examination; and counseling.[1] Individuals are notified of the screening results by mail. In the report, patients are informed of their blood sugar level. They are also informed of "normal" and "at risk" levels of blood sugar, which

[1]Screening is distinct from diagnostic testing, which is performed in response to symptoms or signs of disease. The purpose of screening is to identify disease in asymptomatic individuals.

are “under 111” and “111-120”, respectively.[2] Other measures are reported similarly. In addition to the results of each individual measure, individuals are also informed of whether their overall results classify them as “normal”, “risk group”, or “disease suspected”. If one’s blood sugar falls between 111 and 120, this increases the unconditional probability of receiving a “risk group” notification relative to a blood sugar level less than 111. If blood sugar exceeds 120, this increases the unconditional probability of being classified as “disease suspected” (and inherently reduces the probability of being classified “risk group”). The actual notification depends on the individual’s other measures.[3] Individuals classified as “disease suspected” are eligible for and encouraged to undergo a secondary examination.

3.3 Data

This study uses a merged dataset combining administrative data from NHI and NHSP. The sample consists of males born in even-numbered years who participated in general screening in 2002. The data spans 2001-2006 and contains information on gender, age, insurance contribution, screening results for individual measures including blood sugar (measured in mg/dL), the overall health classification, whether an individual undertook a secondary examination, annual medical expenditures, annual hospital days, annual outpatient days, and mortality.

Our main explanatory variable is the blood sugar level in 2002 (baseline). Our key outcomes of interest are cumulative mortality through 2006, 2006 medical expenditures, and 2006 hospital days. We are also interested in whether individuals undertook a secondary

[2]Blood sugar is measured in units of mg/dL. Levels of blood sugar exceeding 120 is associated with diabetes.

[3]Incorporating the other measures, such as blood pressure, is beyond the scope of this paper. However, future work will incorporate this information to isolate the impact of the diabetes specific risk vs. overall health risk. It would also allow comparison of responses to different diseases for which people may have varying degrees of responsiveness to information.

examination, outpatient days in the year of screening, and whether individuals underwent diabetes screening in the next eligible period. Table 3.1 displays summary statistics of the variables of interest by overall health classification.

3.4 Empirical Framework

The objective of this paper is to assess the impact of receiving different overall health classifications on the health and behavior of patients.[4] To do so, we conduct a regression discontinuity analysis at the 111 and 121 blood sugar levels where the probability of being notified of a particular health status is discontinuous. Specifically, the aim is to compare outcomes across individuals who are effectively identical but for receiving different health status notifications.[5]

The corresponding regression model we estimate is:

$$\text{outcome} = \beta \mathbb{1}\{S \geq \tau\} + f(S) + \epsilon, \tag{3.1}$$

where S is the blood sugar level, $f(S)$ is a function of the blood sugar level, and τ is the relevant cutoff (111 or 121).

In implementing the regression discontinuity design, an important consideration is the modeling of $f(S)$. One approach is to model it parametrically through linear, quadratic, or higher order polynomials that are allowed to differ on each side of the cutoff. The other approach, which we follow here, is to estimate the discontinuity nonparametrically, which we implement by local linear regression with a rectangular kernel.[6] Our preferred estimates

---

[4] In the case of medically related actions, this is likely to be the joint behavior of doctors and their patients.

[5] Because of the nature of the empirical design and data limitations, we are *not* assessing the impact of screening vs. not screening. Rather, we are assessing the impact of the information obtained from screening.

[6] As noted in Lee and Lemieux (2010), the choice of kernel typically has little impact and while a triangular

are based on a bandwidth of 5 mg/dL, in order to reduce bias by staying close to the cutoff while still maintaining enough precision.[7]

A critical assumption to our identification strategy is that individuals just below a threshold are indeed comparable to individuals just above a threshold. One potential threat to this assumption is if individuals are able to precisely sort around the threshold (Lee (2008)). It is not likely that individuals are able to precisely manipulate their blood sugar level. However, other features such as measurement may be just as problematic if it led to nonrandom sorting of scores according to some unobservable characteristic if this characteristic were also correlated with our outcomes of interest. An example of this would be if hospitals that served certain types of patients recorded blood glucose levels above 120 mg/dL as 120 mg/dL.

If our original assumption holds, then an implication is that the density of scores should be continuous around the threshold (indeed, everywhere). Figure 3.1 displays the density of scores for the majority of our sample.[8] Note that there are striking discontinuities in the density at 110 and 120, just below the relevant thresholds. One possible explanation is that there was rounding down (e.g. a score of 114 being recorded as 110), but this would not fully explain the pattern observed (i.e. the reductions in density after 110 and 120).

As discussed in Urquiola and Verhoogen (2009), stacking alone may not violate the regression discontinuity assumptions since violation arises from the interaction of the stacking and the endogenous sorting of individuals. Thus, the more fundamental question for our identification strategy is whether the distribution of predetermined characteristics is identical on each side of the threshold. Figure 3.2 displays the baseline measures of our outcomes of interest (for which there is data) as a function of blood sugar level. For the most part,

kernel is boundary optimal, a more transparent way of putting more weight on observations close to the cutoff is to reestimate a rectangular kernel based model using a smaller bandwidth.

[7]Our results are not sensitive to this choice.

[8]Measures outside of [50,150] were removed for visual reasons.

predetermined characteristics appear to be balanced around each threshold.

To partially address the stacking in the density, we conduct "donut" regression discontinuity analyses by omitting observations with blood sugar levels of 110 and 120. For transparency, in our figures we show the mean estimates of our outcomes at all blood sugar levels, including 110 and 120.

3.5 Results

3.5.1 Classification

Figure 3.3 displays the probability of being classified as "normal" as a function of blood sugar measured at baseline. There is a discrete drop in the probability of "normal" status at 111. The fuzziness arises from the other measures which can also affect status. There is no corresponding impact at 121.

Figure 3.4 displays the probability of being classified as "risk group" as a function of baseline blood sugar. There is a discrete increase in the probability of being "risk group" at 111. At 121, there is a discrete decrease in the probability of being "risk group". This is due to the fact that these individuals are informed that they likely have the disease. This is shown explicitly in Figure 3.5 which displays the probability of "disease suspected" as a function of baseline blood sugar.

Table 3.2 presents the corresponding estimates and illustrates the change in information at each threshold. At the 111 threshold, there is a 15 percentage point drop in the probability of "normal" status and a complementary 15 percentage point increase in the probability of "risk group" status. At the 121 threshold, there is a 33 percentage point drop in the probability of "risk group" status and complementary increase of the same magnitude in the probability of "disease suspected" status. *Thus, the 111 threshold captures the marginal* *impact of "risk group" vs "normal" while the 121 threshold captures the marginal impact of "disease suspected" vs. "risk group".*

3.5.2 Future Mortality and Medical Outcomes

The motivation behind screening, particularly for diabetes, is in order for individuals with chronic conditions to manage their disease and limit the occurrence of future preventable negative health shocks and complications. We seek to assess the impact of screening information on such outcomes in this section. First, we assess mortality. This is a useful measure in part because it is an objective, relevant measure of health. However, it is also very important to know in order to assess the extent to which the other outcomes we assess may be affected by survivor bias. Figure 3.6 shows the cumulative mortality over the five years from the time of screening as a function of baseline blood sugar. There is no apparent impact.

We also consider two additional outcomes, annual medical expenses and hospital stays. As Figures 3.7-3.8 and Table 3.3 indicate, it appears that there are no statistically significant impacts of information status on future health outcomes, as measured by annual medical expenses and hospital days five years after screening. It appears that there may be a negative impact on expenses due to "disease suspected", as the precision of our estimates only allow us to rule out decreases larger than 270 USD/year in absolute value.

3.5.3 Responsiveness

In light of the above null results on outcomes, one possibility is that the information led to changes in behavior, but that these changes in behavior led to no discernible impacts on outcomes. A second possibility is that this information led to no observable changes in behavior either. This section addresses some of these possibilities.

Individuals who were informed that they were "disease suspected" were encouraged to undergo a secondary examination. Figure 3.9 depicts the probability that this happened vs. baseline blood sugar. Consistent with the pattern of "disease suspected", it appears that just over half who were informed followed up for a retest.

However, beyond this follow up, there appears to be no discernible impacts on doctor visits within the year of screening, as shown in Figure 3.10 and Table 3.4. This may not be surprising, given that roughly 90% of the population has at least one doctor's visit during the year and that it may be possible for patients to address their concerns with their physician at a routine exam.

Another explanation for no apparent impact is that treatment for diabetes may be costly, both in terms of changes in behavior as well as potential medical expenses. Thus, what may happen is that individuals wait until their next screening to reassess. To address this, we look at the probability of receiving a blood sugar test in the next screening period, within two to three years from the first year of eligibility. While there appears to be a positive impact on future screening due to "disease suspected" status, this is not statistically different from zero. Interestingly the probability of a future test is decreasing in blood sugar. One potential explanation is that individuals with higher blood sugar are more likely to get diagnosed for diabetes outside of a screening, and so are less likely to need screening in the future.

Responses by Insurance Contribution

In this section, we would like to explore whether responses to information differ by education. This is interesting for many reasons, but particularly because education may be an important factor in processing information. Because the data does not include information on education, we use insurance contribution (which is directly proportional to income) as a proxy. The outcome we consider is revisits, mainly because this is the only behavioral response that was statistically significant in the previous section. We focus on the 121 threshold, which is the relevant threshold for this outcome. Table 3.5 summarizes our findings. While there is no particular pattern in the changes in information at 121, there is a fairly clear indication that the top quintile has a lower response to the information than the other quintiles, with only 42% responding with a revisit relative to the 60% of other quintiles. There are several possible explanations. One is that the more responsive among the richer are more educated individuals who are more likely to have already been diagnosed. This explanation is not likely given that there are no substantial differences in information/diagnosis between the top and other quintiles. Another potential explanation is that educated individuals respond less to the (redundant) diagnostic label because they have already largely incorporated this information from the blood sugar measure itself.

3.6 Conclusion and Future Work

This paper studies the impact of information from screening on health outcomes and behavior. We find that encouragement and eligibility for secondary examinations due to a "disease suspected" classification leads to follow-up rates of greater than 50%. However, we find few impacts otherwise, including short and medium run medical activity and longer run health outcomes. On one hand, these results may not be surprising given the relatively low risk of near-term complications at the thresholds we study and the high cost of managing diabetes in terms of changes in behavior. On the other hand, it is surprising to see no evidence of changes in medical activity, given the medical relevance of the thresholds and the ability of providers to influence this action. In either case, our findings suggest the need for futher study of such medically relevant thresholds in order to improve their effectiveness regarding both efficacy as well as cost considerations.

Related to above, there are several avenues of future research interest. Immediate next steps include evaluating the other screening measures such as blood pressure and γ-GTP. This would also enable isolation of disease specific responses from responses to the general classification. Given the chronic nature of diabetes, incorporating a more dynamic analysis may prove fruitful. Lastly, the measures studied here are by no means comprehensive, but are mainly limited due to data considerations. If such data could be gathered, either in this or other contexts, both data on behavior like smoking or diet as well as disease specific medical expenditures would enable more precise and comprehensive analyses.

Table 3.1: Summary Statistics by Blood Sugar Level

Baseline Blood Sugar	**[101,109]**	**[111,119]**	**[121,129]**
# Obs	*86,565*	*38,303*	*13,880*
Normal	0.26	0.07	0.03
	(0.44)	(0.25)	(0.16)
Risk Group	0.33	0.47	0.12
	(0.47)	(0.50)	(0.33)
Disease Suspected	0.41	0.46	0.85
	(0.49)	(0.50)	(0.35)
Insurance	38,362	37,132	37,346
Contribution	(24,941)	(24,461)	(25,314)
Baseline Medical	477	516	590
Expenditures	(822)	(902)	(1,029)
Baseline	0.78	0.84	0.98
Hospital Days	(5.48)	(5.54)	(6.34)
Baseline	14.4	15.3	16.4
Outpatient Days	(17.2)	(18.2)	(19.6)
5 Year Cumulative	0.03	0.04	0.04
Mortality	(0.18)	(0.19)	(0.20)
Medical Expenditures	1,208	1,335	1,504
5 Years Later	(2,771)	(2,914)	(2,863)
Hospital Days	2.22	2.45	2.87
5 Years Later	(14.10)	(15.32)	(17.69)
Secondary Exam	0.22	0.26	0.48
	(0.42)	(0.44)	(0.50)
Outpatient Days	15.8	16.7	17.9
Screening Year	(18.6)	(19.5)	(20.4)
DM Screen	0.77	0.75	0.73
Next Period	(0.42)	(0.43)	(0.44)

Notes: Sample consists of male individuals born in even-numbered years and who participated in general screening in 2002. Categories are based on baseline blood sugar levels. Individuals with blood sugar levels of 110 and 120 were omitted from the analysis as discussed in the text. See text for definitions of variables.

Table 3.2: Impact of Baseline Blood Sugar Level on Notified Status

	Normal	**At-Risk**	**Disease**
Blood Sugar ≥ 111	-0.15**	0.15**	-0.01
	(0.01)	(0.01)	(0.01)
Blood Sugar ≥ 121	-0.01	-0.33**	0.34**
	(0.01)	(0.01)	(0.02)

Notes: Each cell presents estimates of β from local linear regression of Equation (3.1). The running variable is baseline blood sugar level. Rectangular kernel. Robust standard errors in parentheses. ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

Table 3.3: Impact of Baseline Blood Sugar Level on Outcomes 5 Years After Screening

	Mortality	**Med Exp**	**Hosp Days**
Blood Sugar ≥ 111	0.00	119.21+	0.32
	(0.00)	(62.03)	(0.33)
Blood Sugar ≥ 121	-0.01	-88.42	-0.25
	(0.01)	(94.49)	(0.53)

Notes: Each cell presents estimates of β from local linear regression of Equation (3.1). The running variable is baseline blood sugar level. Rectangular kernel. Robust standard errors in parentheses. ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

Table 3.4: Impact of Baseline Blood Sugar Level on Medical Behavior

	Retest	**OPT Days**	**DM Screen**
Blood Sugar ≥ 111	0.00	0.36	0.01
	(0.01)	(0.42)	(0.01)
Blood Sugar ≥ 121	0.19**	-0.79	0.02
	(0.02)	(0.71)	(0.01)

Notes: Each cell presents estimates of β from local linear regression of Equation (3.1). The running variable is baseline blood sugar level. Rectangular kernel. Robust standard errors in parentheses. ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

Table 3.5: Revisit Percentage by Income

	Disease	Retest	Retest/Disease
Top Quintile	0.36**	0.15**	0.42**
	(0.03)	(0.03)	(0.08)
2nd Quintile	0.27**	0.17**	0.65**
	(0.03)	(0.03)	(0.11)
3rd Quintile	0.39**	0.24**	0.61**
	(0.04)	(0.04)	(0.07)
4th Quintile	0.34**	0.19**	0.57**
	(0.03)	(0.03)	(0.08)
Bottom Quintile	0.32**	0.19**	0.60**
	(0.03)	(0.03)	(0.09)

Notes: Columns 1 and 2 present estimates of β from local linear regression of Equation (3.1) at the 121 threshold. Column 3 is Column 2 divided by Column 1. Each row is a different income quintile. The running variable is baseline blood sugar level. Rectangular kernel. Robust standard errors in parentheses. ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

Figure 3.1: Histogram of Baseline Blood Sugar

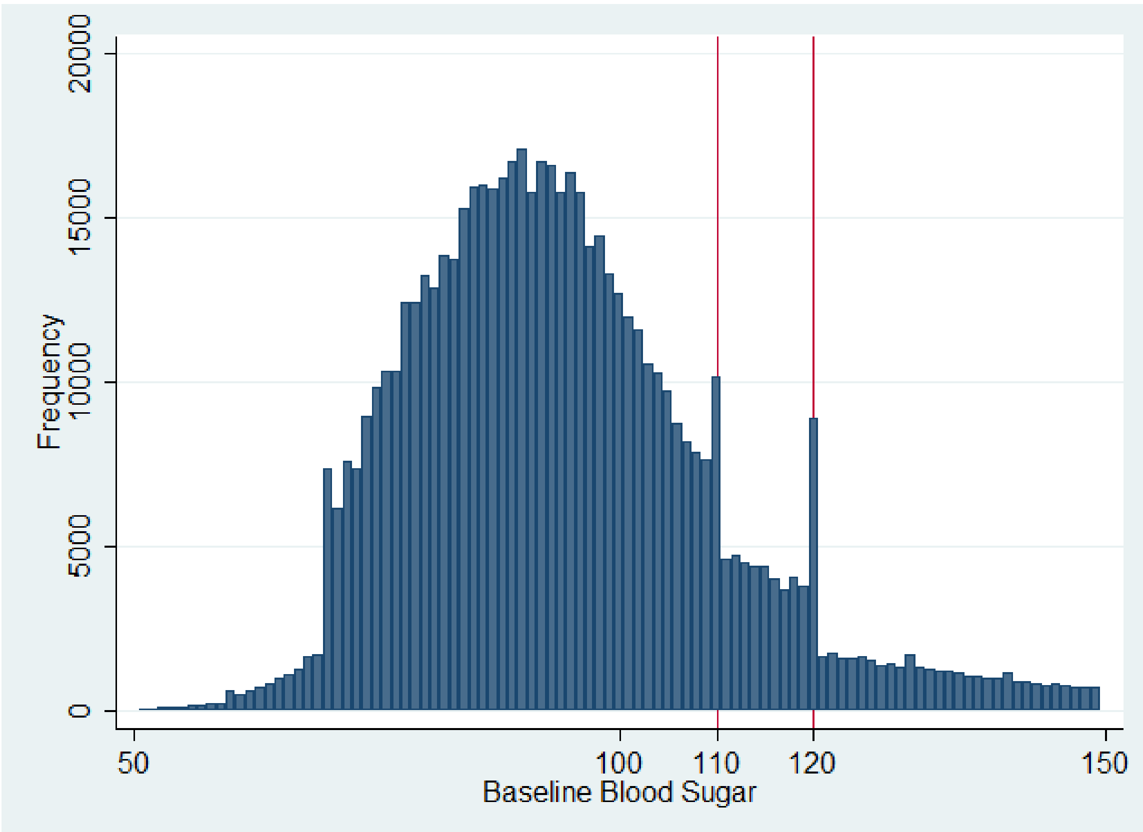
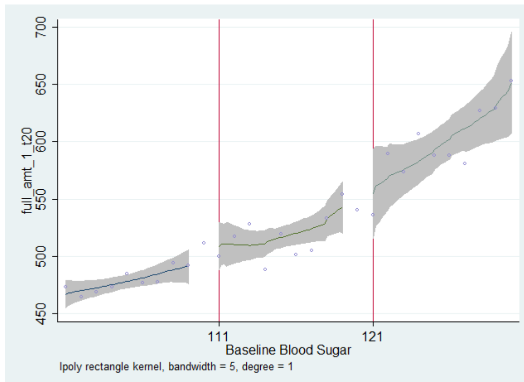
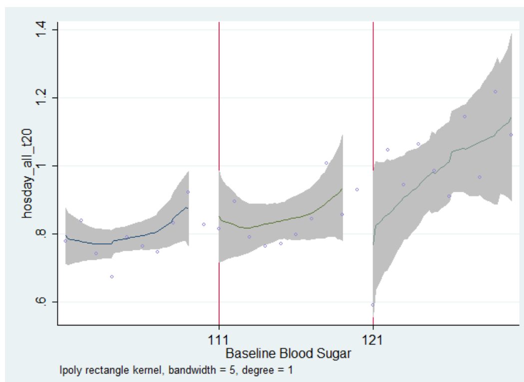


Figure 3.2: Baseline Characteristics

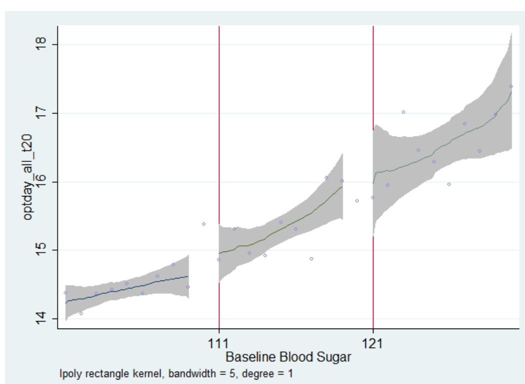
(a) Annual Medical Expenses



(b) Annual Hospital Days

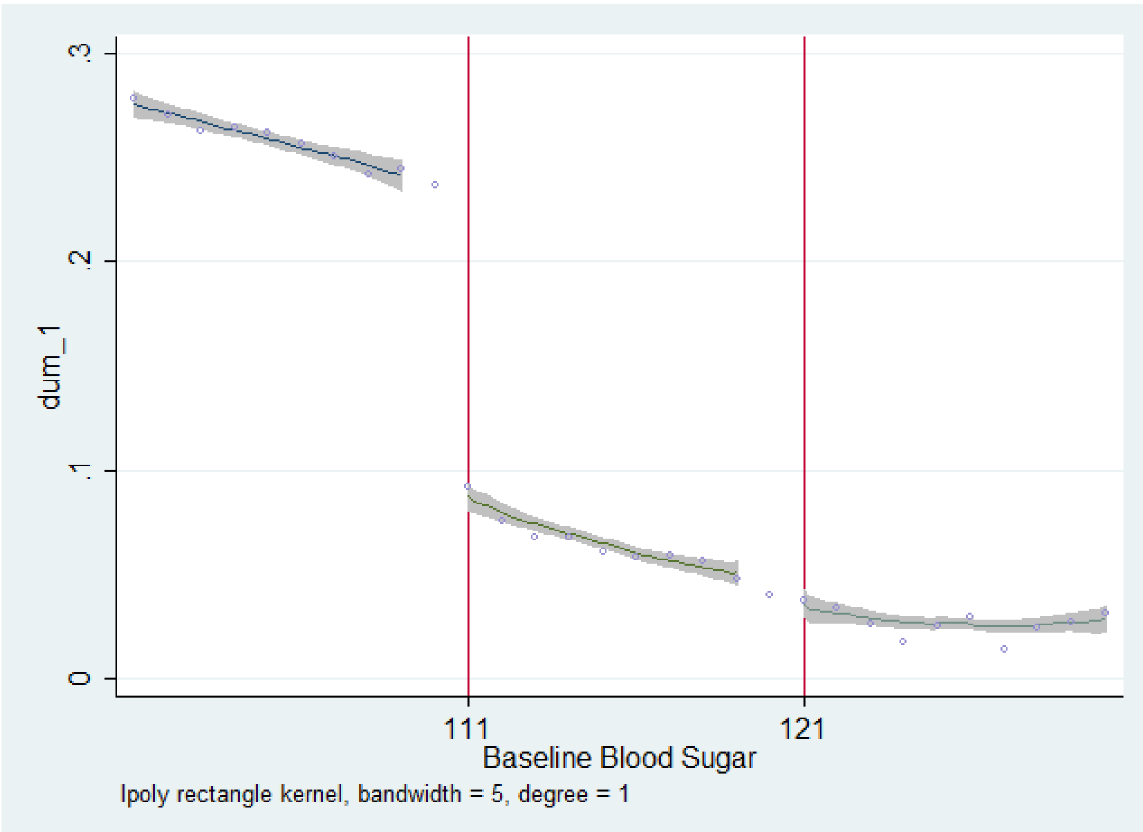


(c) Annual Outpatient Days



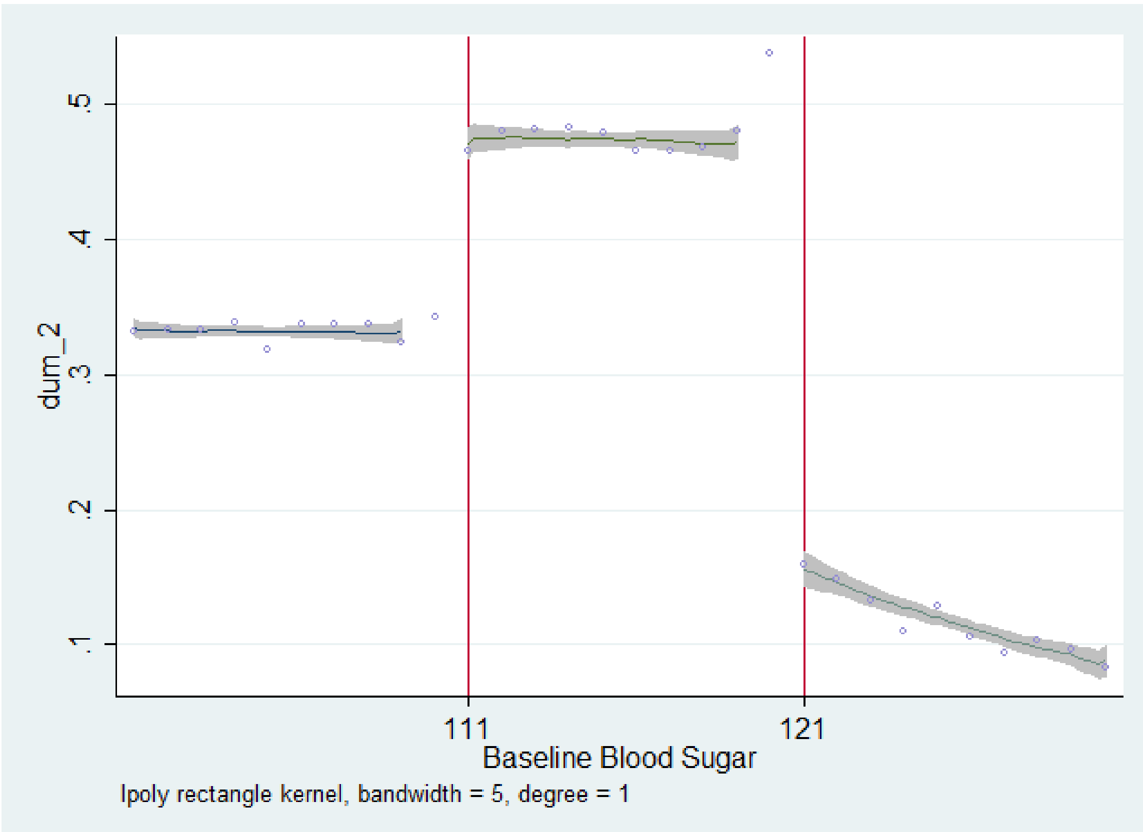
Notes: The running variable is baseline blood sugar level. The open circles plot the mean of the dependent variable at each unit. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 5 mg/dL. The shaded regions are 95 percent confidence intervals.

Figure 3.3: "Normal" Status



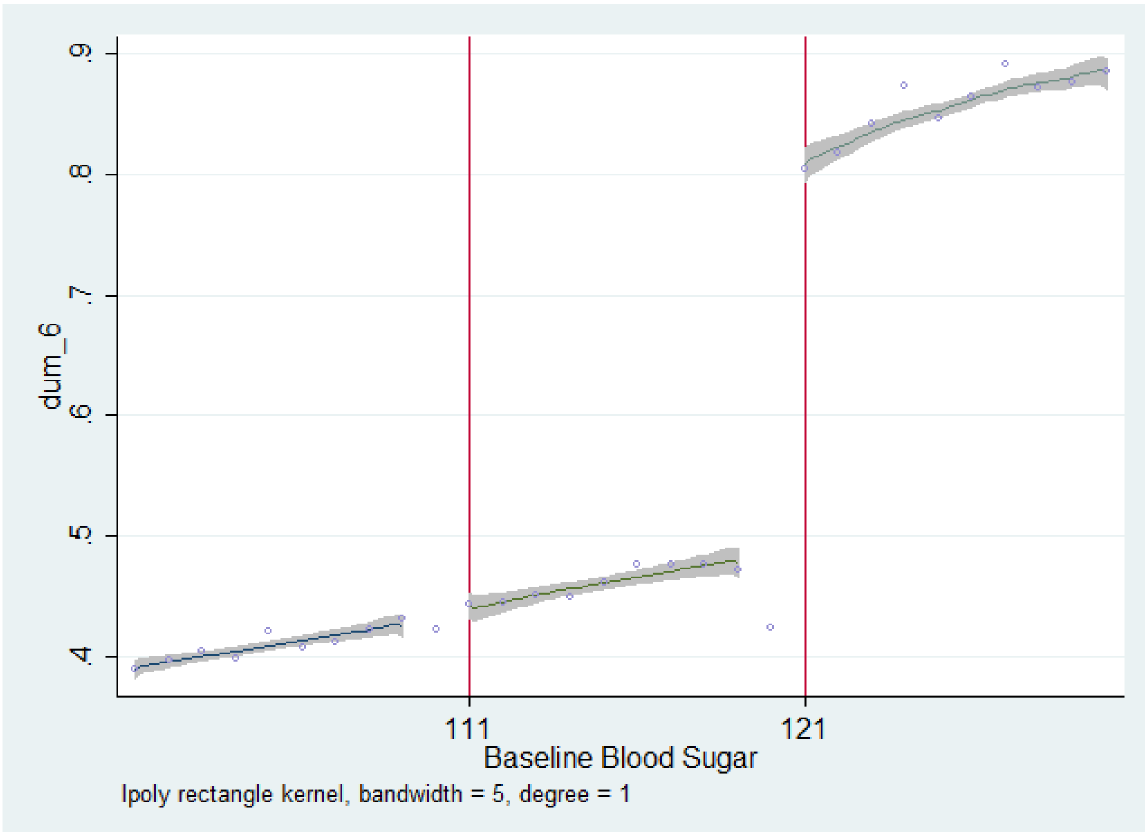
Notes: The running variable is baseline blood sugar level. The open circles plot the mean of the dependent variable at each unit. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 5 mg/dL. The shaded regions are 95 percent confidence intervals.

Figure 3.4: "Risk Group" Status



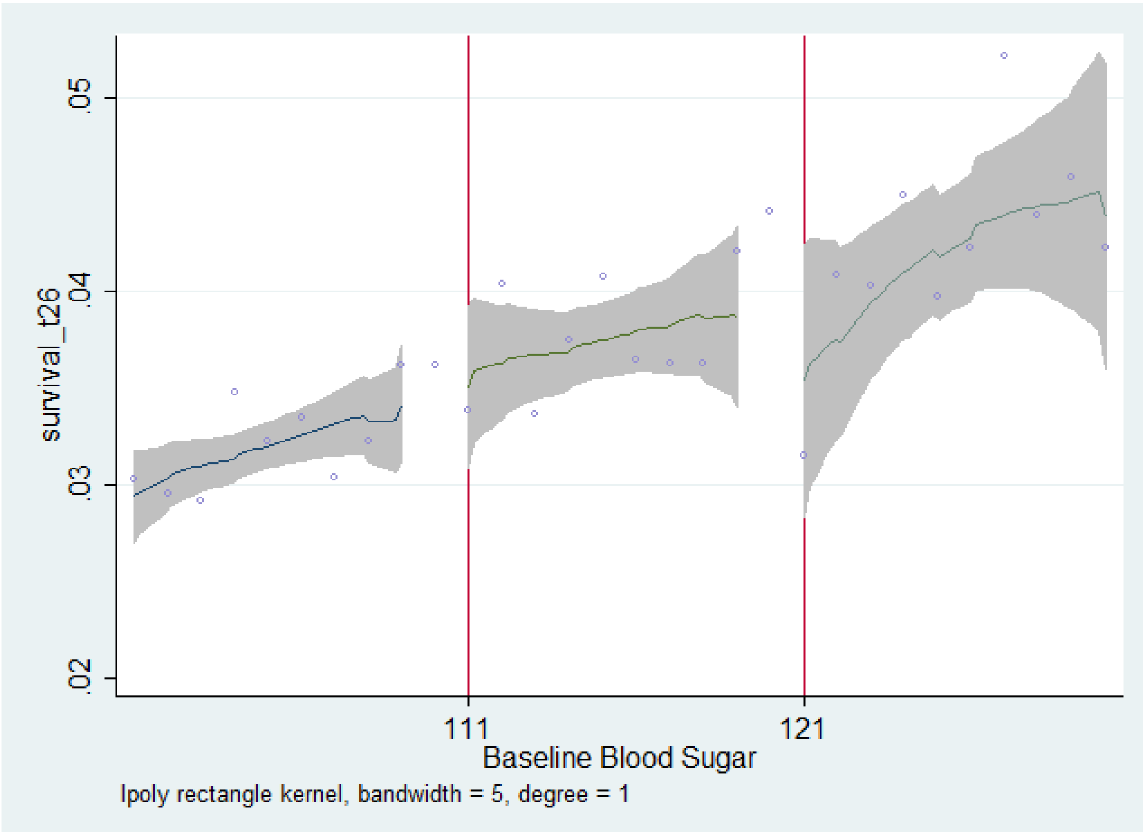
Notes: The running variable is baseline blood sugar level. The open circles plot the mean of the dependent variable at each unit. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 5 mg/dL. The shaded regions are 95 percent confidence intervals.

Figure 3.5: "Disease Suspected" Status



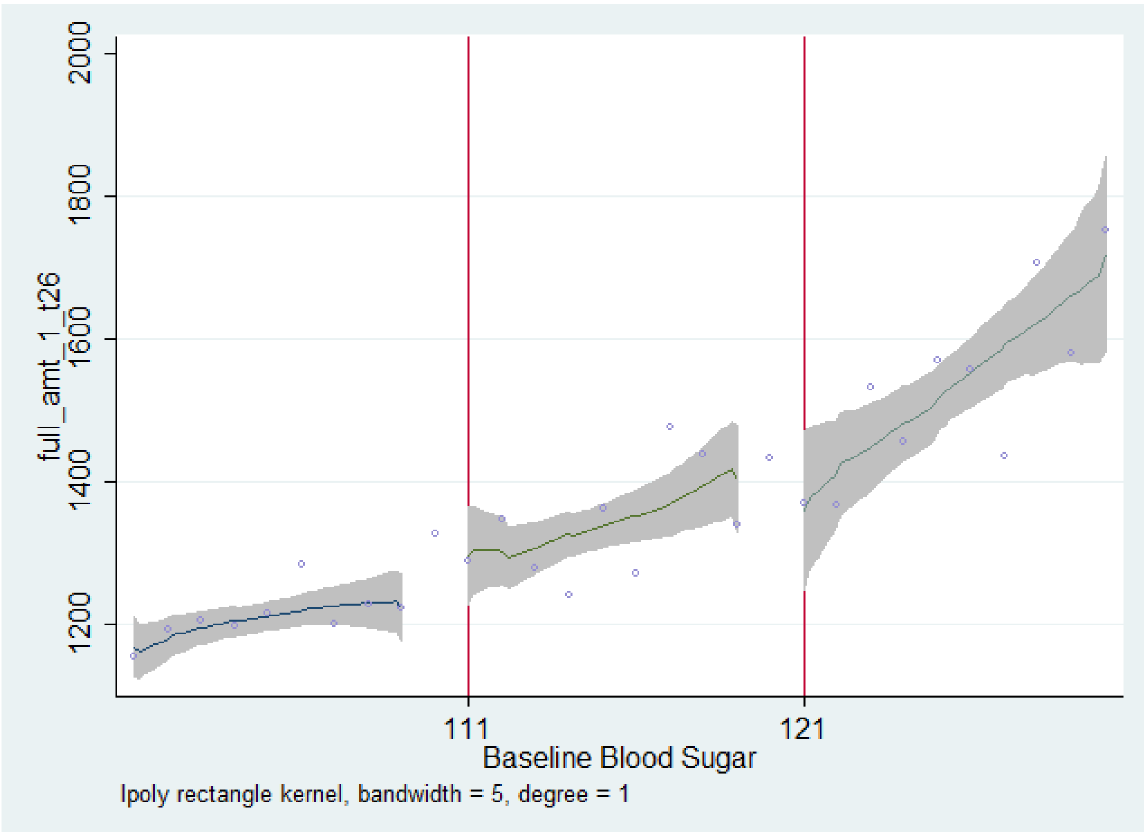
Notes: The running variable is baseline blood sugar level. The open circles plot the mean of the dependent variable at each unit. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 5 mg/dL. The shaded regions are 95 percent confidence intervals.

Figure 3.6: Cumulative Mortality Through 5 Years After Screening



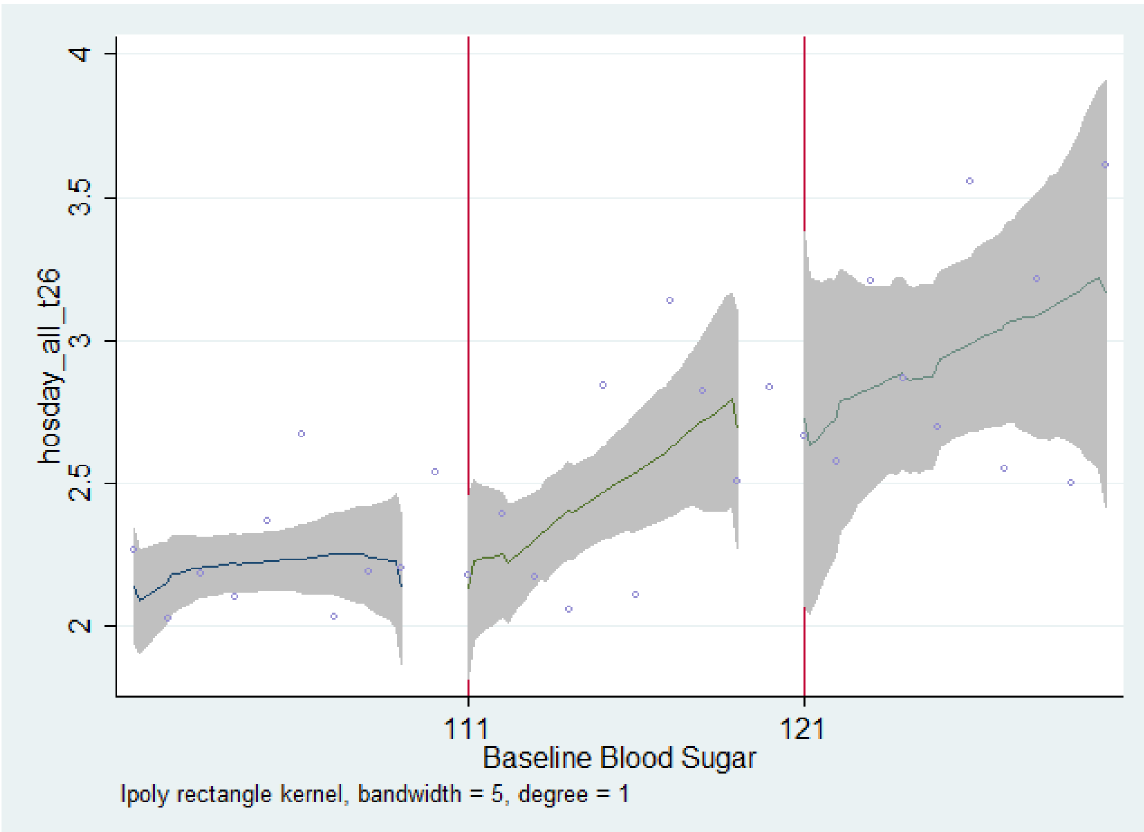
Notes: The running variable is baseline blood sugar level. The open circles plot the mean of the dependent variable at each unit. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 5 mg/dL. The shaded regions are 95 percent confidence intervals.

Figure 3.7: Annual Medical Expenses 5 Years Later



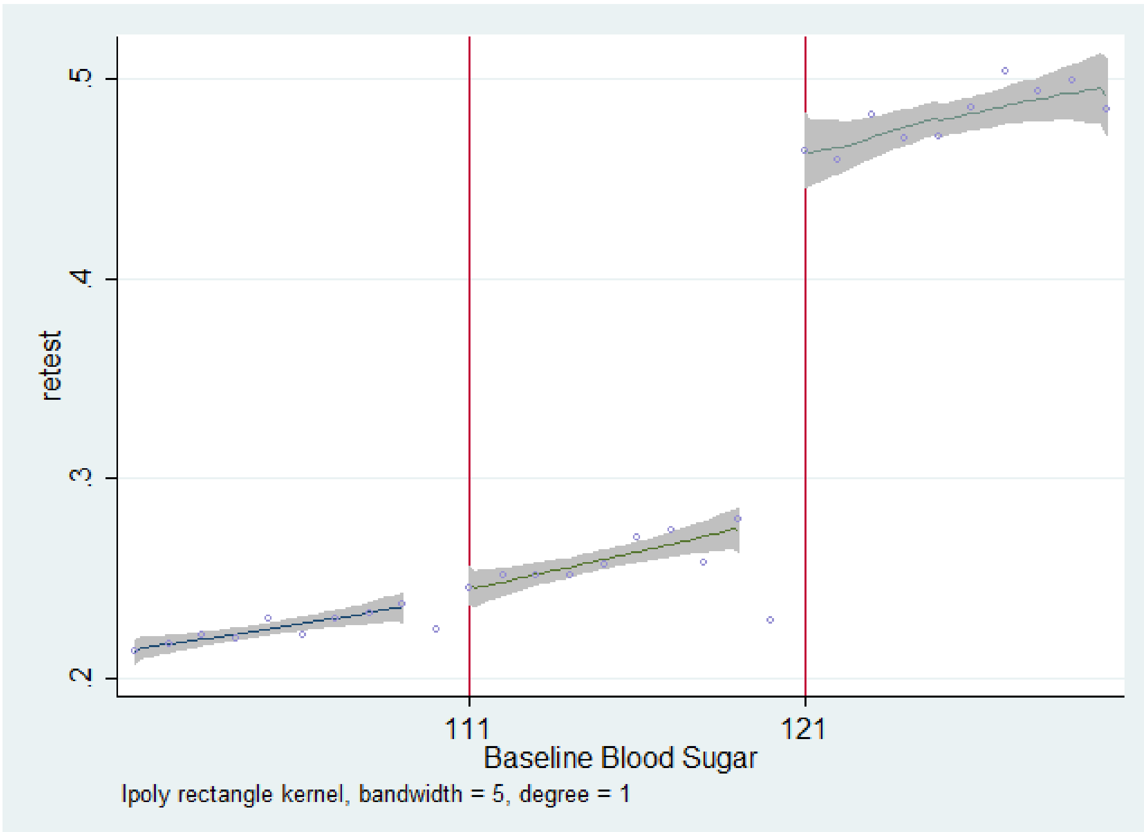
Notes: The running variable is baseline blood sugar level. The open circles plot the mean of the dependent variable at each unit. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 5 mg/dL. The shaded regions are 95 percent confidence intervals.

Figure 3.8: Annual Hospital Days 5 Years Later



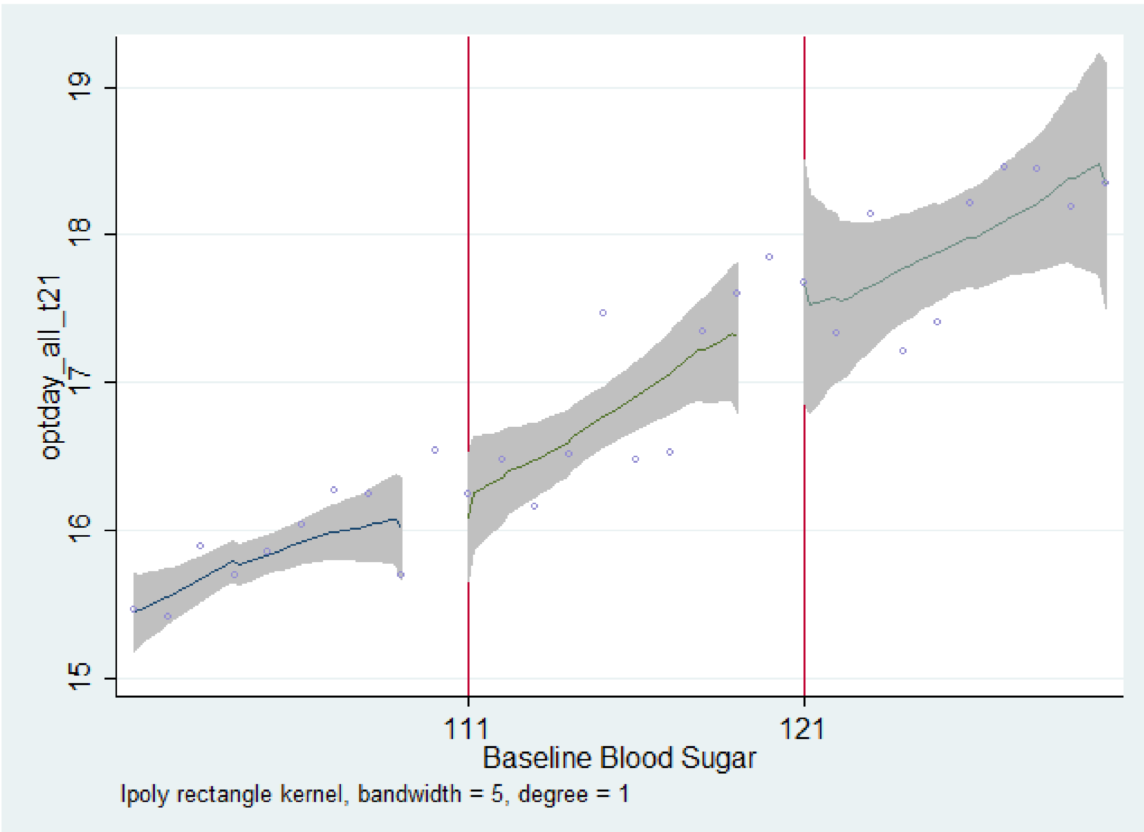
Notes: The running variable is baseline blood sugar level. The open circles plot the mean of the dependent variable at each unit. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 5 mg/dL. The shaded regions are 95 percent confidence intervals.

Figure 3.9: Clinic Revisit



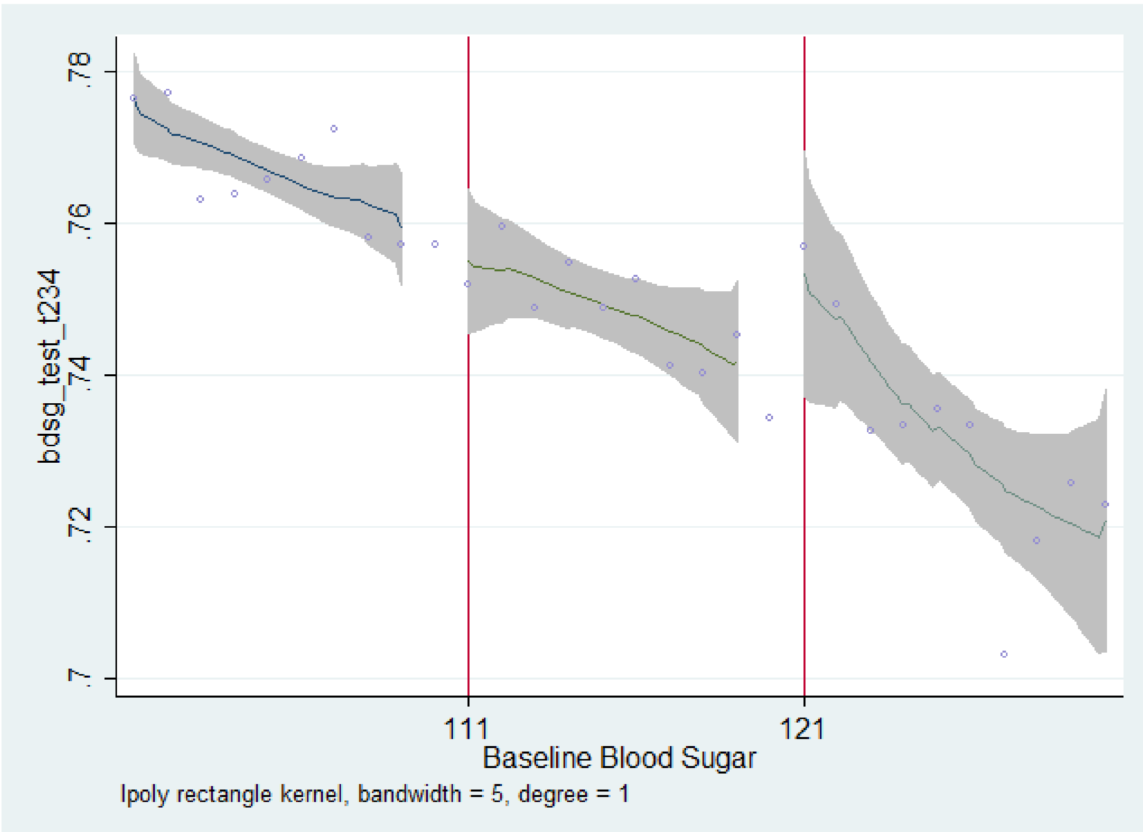
Notes: The running variable is baseline blood sugar level. The open circles plot the mean of the dependent variable at each unit. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 5 mg/dL. The shaded regions are 95 percent confidence intervals.

Figure 3.10: Outpatient Days



Notes: The running variable is baseline blood sugar level. The open circles plot the mean of the dependent variable at each unit. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 5 mg/dL. The shaded regions are 95 percent confidence intervals.

Figure 3.11: Take Blood Sugar Test at Next Screening Opportunity



Notes: The running variable is baseline blood sugar level. The open circles plot the mean of the dependent variable at each unit. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 5 mg/dL. The shaded regions are 95 percent confidence intervals.

Bibliography

Bass, DM and LS Noelker. 1987. “The Influence of Family Caregivers on Elder Use of In-Home Services : An Expanded Conceptual Framework.” *Journal of Health and Social Behavior* 28 (2):184–196.

Bitler, Marianne P. and Christopher S. Carpenter. 2011. “Insurance Mandates and Mammography.” Working Paper 16669, National Bureau of Economic Research.

Busch, Susan H. and Colleen L. Barry. 2008. “New Evidence on the Effects of State Mental Health Mandates.” *Inquiry* 45 (3):308–322.

CBO. 2007. *The Long-Term Outlook for Health Care Spending*. Congressional Budget Office.

CDC. 2011. *National Diabetes Fact Sheet*. Centers for Disease Control.

Charles, KK and P Sevak. 2005. “Can Family Caregiving Substitute for Nursing Home Care?” *Journal of Health Economics* 24 (6):1174–1190.

Christianson, JB. 1988. “The Evaluation of the National Long-Term Care Demonstration .6. The Effect of Channeling on Informal Caregiving.” *Health Services Research* 23 (1):99–117.

Colombo, F., A. Llena-Nozal, J. Mercier, and F. Tjadens. 2011. *Help Wanted? Providing and Paying for Long-Term Care*. OECD Health Policy Studies. OECD Publishing.

Cutler, David M. and Adriana Lleras-Muney. 2006. “Education and Health: Evaluating Theories and Evidence.” Working Paper 12352, National Bureau of Economic Research.

Engelhardt, Gary V and Nadia Greenhalgh-Stanley. 2010. “Home Health Care and the Housing and Living Arrangements of the Elderly.” *Journal of Urban Economics* 67 (2):226–238.

GAO. 2005. *Long-Term Care Financing: Growing Demand and Cost of Services are Straining Federal and State Budgets*. U.S. Government Accountability Office.

Golberstein, E, DC Grabowski, KM Langa, and ME Chernew. 2009. “Effect of Medicare Home Health Care Payment on Informal Care.” *Inquiry* 46 (1):58–71.

Grossman, Michael. 2006. “Education and Nonmarket Outcomes.” Elsevier, 577 – 633.

Gruber, J. 1994. "The Incidence of Mandated Maternity Benefits." *American Economic Review* 84 (3):622–641.

Imbens, Guido and Karthik Kalyanaraman. 2009. "Optimal Bandwidth Choice for the Regression Discontinuity Estimator." Working Paper 14726, National Bureau of Economic Research.

Kahn, ME. 1999. "Diabetic Risk Taking: The Role of Information, Education and Medication." *Journal of Risk and Uncertainty* 18 (2):147–164.

Klick, Jonathan and Thomas Stratmann. 2007. "Diabetes Treatments and Moral Hazard." *Journal of Law & Economics* 50 (3):519–538.

Lange, Fabian. 2011. "The Role of Education in Complex Health Decisions: Evidence from Cancer Screening." *Journal of Health Economics* 30 (1):43–54.

Lee, David S. 2008. "Randomized Experiments from Non-Random Selection in US House Elections." *Journal of Econometrics* 142 (2):675–697.

Lee, David S and Thomas Lemieux. 2010. "Regression Discontinuity Designs in Economics." *Journal of Econometric Literature* 48 (2):281–355.

Li, Rui, Ping Zhang, Lawrence Barker, and DeKeely Hartsfield. 2010. "Impact of State Mandatory Insurance Coverage on the Use of Diabetes Preventive Care." *BMC Health Services Research* 10.

Lleras-Muney, A. 2005. "The Relationship Between Education and Adult Mortality in the United States." *Review of Economic Studies* 72 (1):189–221.

Lo Sasso, AT and RW Johnson. 2002. "Does Informal Care from Adult Children Reduce Nursing Home Admissions for the Elderly?" *Inquiry* 39 (3):279–297.

McKnight, R. 2006. "Home Care Reimbursement, Long-Term Care Utilization, and Health Outcomes." *Journal of Public Economics* 90 (1-2):293–323.

Orsini, Chiara. 2010. "Changing the Way the Elderly Live: Evidence from the Home Health care Market in the United States." *Journal of Public Economics* 94 (1-2):142–152.

Peltzman, S. 1975. "Effects of Automobile Safety Regulation." *Journal of Political Economy* 83 (4):677–725.

Pezzin, LE, P Kemper, and J Reschovsky. 1996. "Does Publicly Provided Home Care Substitute for Family Care? Experimental Evidence with Endogenous Living Arrangements." *Journal of Human Resources* 31 (3):650–676.

Soldo, BJ. 1985. "In-Home Services for the Dependent Elderly: Determinants of Current Use and Implications for Future Demand." *Research on Aging* 7 (2):281–304.

Stabile, Mark, Audrey Laporte, and Peter C. Coyte. 2006. "Household Responses to Public Home Care Programs." *Journal of Health Economics* 25 (4):674–701.

Urquiola, Miguel and Eric Verhoogen. 2009. "Class-Size Caps, Sorting, and the Regression-Discontinuity Design." *American Economic Review* 99 (1):179–215.

USPSTF. 2008. "Screening for Prostate Cancer: US Preventive Services Task Force Recommendation Statement." *Annals of Internal Medicine* 149 (3):185–191.

Van Houtven, CH and EC Norton. 2004. "Informal Care and Health Care Use of Older Adults." *Journal of Health Economics* 23 (6):1159–1180.

Ward, Derek, Amy Drahota, Diane Gal, Martin Severs, and Taraneh P. Dean. 2008. "Care Home Versus Hospital and Own Home Environments for Rehabilitation of Older People." *Cochrane Database of Systematic Reviews* (4).

Wolinsky, FD, RR Mosely, and RM Coe. 1986. "A Cohort Analysis of the Use of Health Services by Elderly Americans." *Journal of Health and Social Behavior* 27 (3):209–219.

Wooldridge, J and J Schore. 1988. "The Evaluation of the National Long-Term Care Demonstration .7. The Effect of Channeling on the Use of Nursing Homes, Hospitals, and Other Medical Services." *Health Services Research* 23 (1):119–127.

Appendix A

Calculation of the Preliminary Score

The preliminary score is calculated from the responses to 52 questions in the eligibility evaluation. A list of these items and possible responses are listed in Table A.1.

The procedure for determining the preliminary score is as follows:

1. Convert responses to point values, according to Table A.1.
2. Sum the points in each category.
3. Based on the category scores and the responses to the 52 items, determine sub-scores for eight service categories: individual hygiene, excretion support, eating, moving, behavior, indirect support, nursing care, and rehabilitation. See Figure A.1 for an illustration of how the eating sub-score is determined.
4. Sum the service sub-scores to arrive at the preliminary score.

We now provide a partial example for calculating the preliminary score. Table A.2 contains a sample set of answers to the eligibility assessment. The category scores for ADL and REH are 16 and 13, respectively. Now follow the eating tree. The first fork depends on the response to “eating” in the ADL category. The response of independent (1) sends us along the left branch. The response to ”brushing teeth” is independent (1), which takes us down the first branch. Since the ADL score is 16, we end up along the right branch for a score of 9.4. We repeat this procedure for the remaining service sub-scores:

- Individual hygiene 5.3
- Excretion support 2.6
- Eating 9.4
- Moving 3.6
- Behavior 0.8
- Indirect 21.7
- Nursing Care 9.7
- Rehabilitation 2.7

These sum to the final score of 55.8.

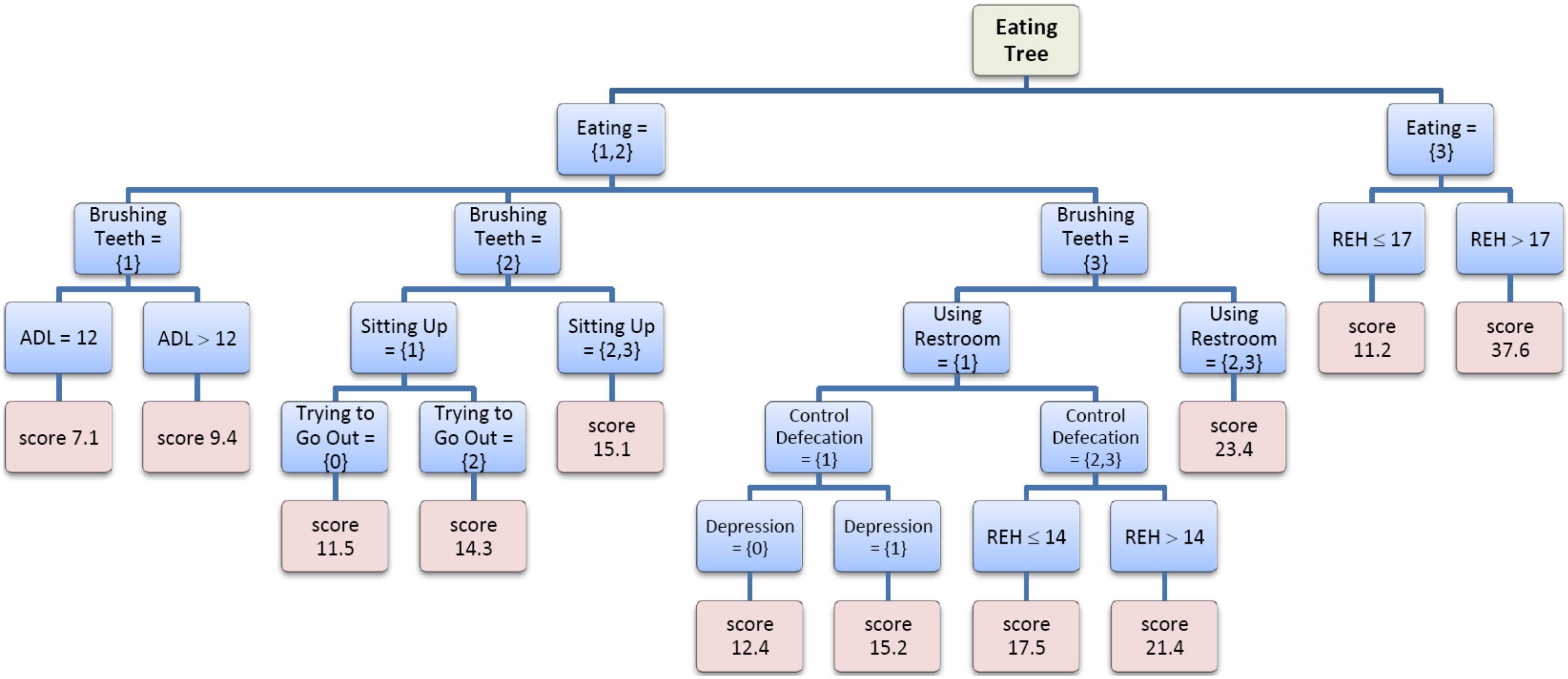
Table A.1: Assessment Questions

Category	**Items**	**Possible Responses (response score)**
Physical Function	Dressing and Undressing, Washing His/Her Face, Brushing His/Her Teeth, Bathing, Eating, Changing Positions, Sitting Up, Moving to Sit Elsewhere, Moving to Another Room, Using the Restroom, Control of Defecation, Control of Urination	Independent (1), Needs Partial Help (2), Needs Complete Help (3)
Cognitive Function	Short Term Memory Disorder, Failing to Perceive Date, Failing to Perceive Location, Failing to Recall Age and Birth Date, Failing to Understand Instructions, Deteriorating Circumstantial Judgement, Communication Disorder	No (0), Yes (1)
Behavior	Hallucination, Delusion, Depression, Sleep Disorder, Resistant to Help, Restless, Gets Lost, Abusive/Aggressive Behavior, Trying to Go Out, Breaks Things, Inappropriate Behavior, Hiding Money/Goods, Dressing Improperly, Lack of Restroom Hygiene	No (0), Yes (1)
Medical Treatment	Tracheotomy, Aspiration, Oxygen Therapy, Bed Sore, Tube Feeding, Cancer Pain Management, Urine Catheter, Fistula Care, Dialysis	No (0), Yes (1)
Rehabilitation	Motor Disturbance (Right/Left Arm/Leg); Joint Disorder (Shoulder, Elbow, Wrist, Hip , Knee, Ankle)	No Disorder/Limitation (1), Partial Disorder/Limitation (2), Complete Disorder/Limitation (3)

Table A.2: Sample Assessment

Category	Item	Response	Score
Activities of Daily Living	Dressing and Undressing	Need Partial Support (NPS)	2
	Washing Face	Fully Self Support (FSS)	1
	Brushing Teeth	Fully Self Support (FSS)	1
	Bathing	Need Full Support (NFS)	3
	Feeding	Fully Self Support (FSS)	1
	Changing Position	Fully Self Support (FSS)	1
	Sitting Up	Fully Self Support (FSS)	1
	Changing Sitting Location	Fully Self Support (FSS)	1
	Ambulation	Need Partial Support (NPS)	2
	Using the Restroom	Fully Self Support (FSS)	1
	Voluntarily Control of Fecal Discharge	Fully Self Support (FSS)	1
	Voluntarily Controlling of Urinary Discharge	Fully Self Support (FSS)	1
			16
Cognitive Function	Short Term Memory Loss	Yes	1
	Disorientation of Date	Yes	1
	Disorientation of Place	Yes	1
	Disorientation of Age and Birth Date	No	0
	Disorientation of Order	Yes	1
	Disorientation of Judgement	Yes	1
	Despair of Communication	No	0
			5
Misbehavior	Illusion	No	0
	Delusion	Yes	1
	Depression	No	0
	Sleep Disorder	Yes	1
	Resistant to Support	No	0
	Being Anxious / Lingering Around	Yes	1
	Being Lost	No	0
	Abusive Language / Aggressive Behavior	No	0
	Trying to Go Out	Yes	1
	Destroys Things	No	0
	Meaningless or Inappropriate Behavior	Yes	1
	Hiding Money / Things	Yes	1
	Inappropriate Clothing	No	0
	Unclean Urination / Defecation	No	0
			6
Nursing	Tracheotomy	No	0
	Aspiration	No	0
	Oxygen Therapy	No	0
	Bed sore	No	0
	Tube Feeding	No	0
	Pain Control of Cancer	No	0
	Care of Urine Catheter	No	0
	Fistula Care	No	0
	Care for Dialysis	No	0
			0
Rehabilitation	Right Arm	No Disorder (ND)	1
	Left Arm	No Disorder (ND)	1
	Right Leg	No Disorder (ND)	1
	Left Leg	No Disorder (ND)	1
	Shoulder	No Limitation	1
	Elbow	No Limitation	1
	Wrist	No Limitation	1
	Hip Joint	Symmetry	3
	Knee Joint	Asymmetry	2
	Ankle	No Limitation	1
			13

Figure A.1: Eating Tree



Appendix B

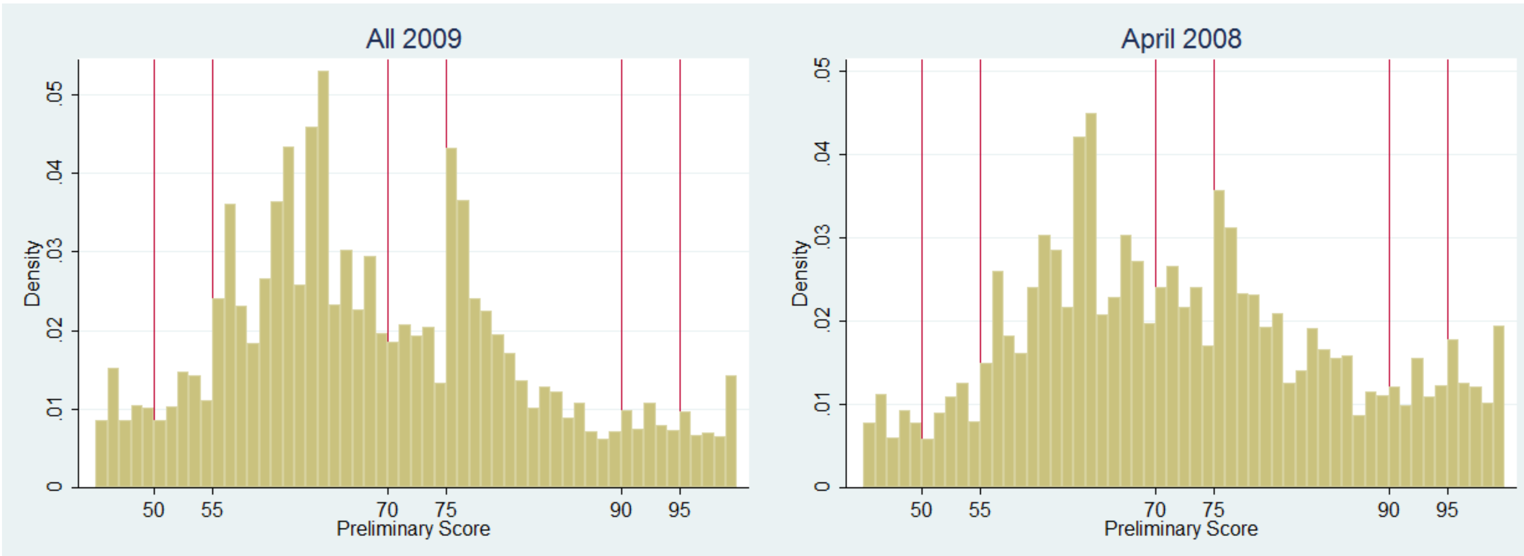
Supplemental Tables and Figures

Table B.1: Description of Reimbursed LTC Services

Category	Benefit	Description
Home	Home Help Visit	A care worker visits the beneficiary's residence and helps in the following: bathing, using the restroom, changing clothes, washing hair, cooking, buying daily necessities, cleaning, and clearing up surroundings.
Home	Home Bathing Visit	A care worker visits the beneficiary's residence to provide bathing services using a bathing device.
Home	Home Nursing Visit	A nurse or dental hygienist visits the beneficiary's residence to provide nursing, treatment assistance, care consultation, or dental hygiene services.
Home	Day and Night Care	Facility stay for less than a day where education or training is provided to the beneficiary for maintenance or improvement of physical and mental function.
Home	Short-Term Respite Care	Short term facility care in order to provide temporary relief for the regular caregiver.
Home	Medical Supplies	Equipment is provided for the support of the beneficiary's daily tasks and physical activities (e.g. bath seat or walker).
Institutional	Elder Care Facility or Group Home	Residence, meals, care, and other conveniences requried for daily function.

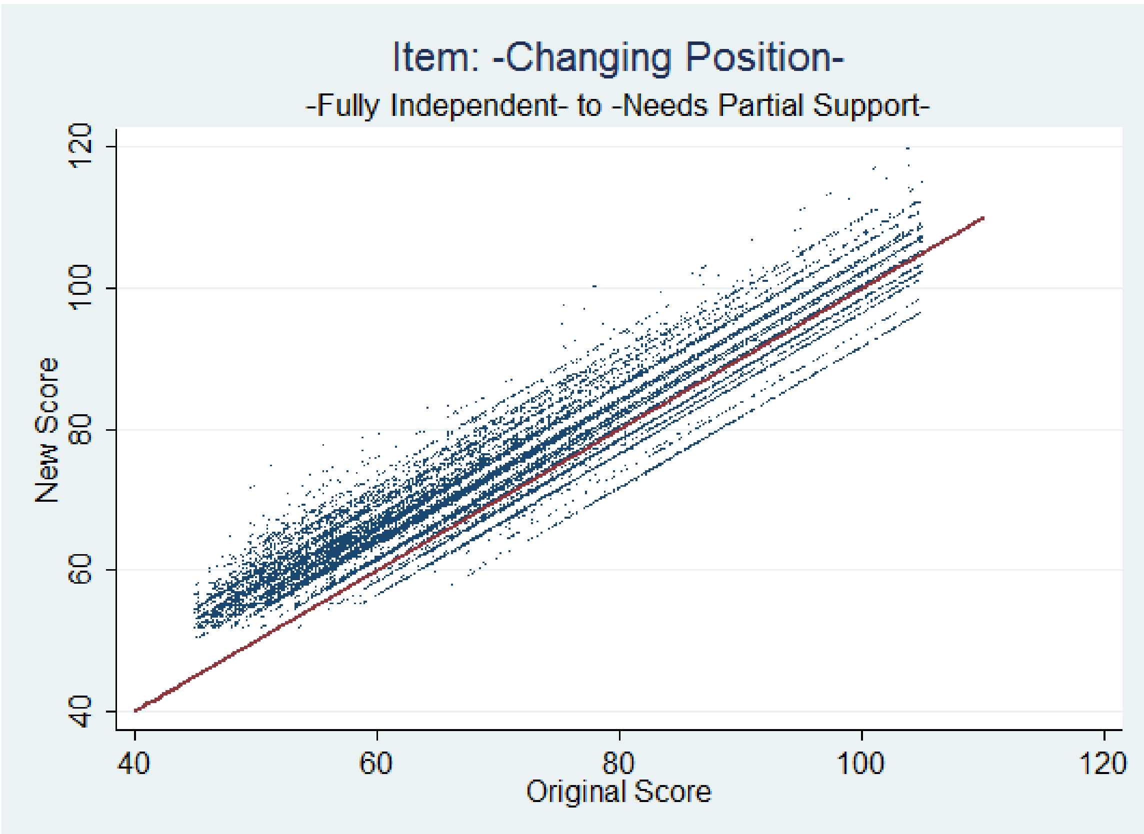
Figure B.1: Density of Scores, 2009 vs. April 2008

(a) (b)



Notes: Preliminary scores in 1 point bins. “All 2009” consists of the preliminary scores from assessments in 2009. “April 2008” consists of the preliminary scores from assessments in April 2008.

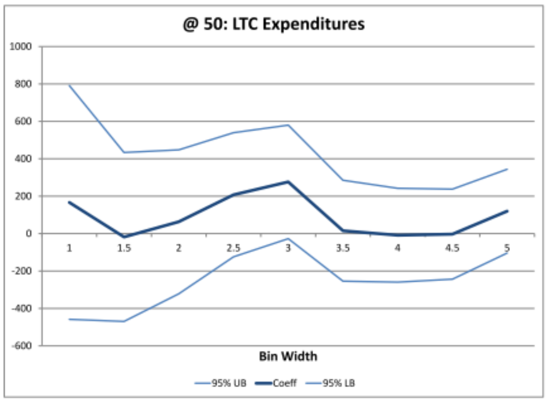
Figure B.2: Score Sensitivity Example



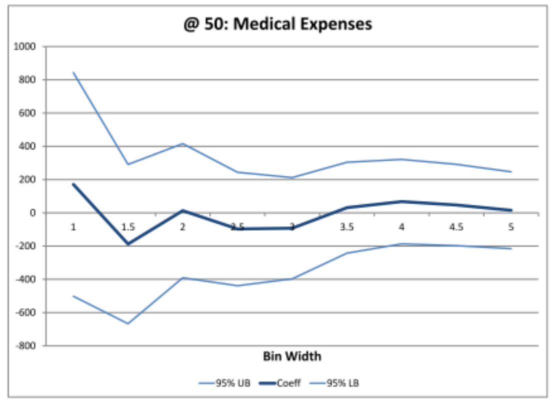
Notes: Sample of individuals whose preliminary scores fall between 45 and 105, with the response "Fully Independent" to the item "Changing Position". The original preliminary score is on the x-axis. The new preliminary score after changing the response from "Fully Independent" to "Needs Partial Support" is on the y-axis. Also graphed in red is the 45 degree line.

Figure B.3: Sensitivity to Bandwidth at 50

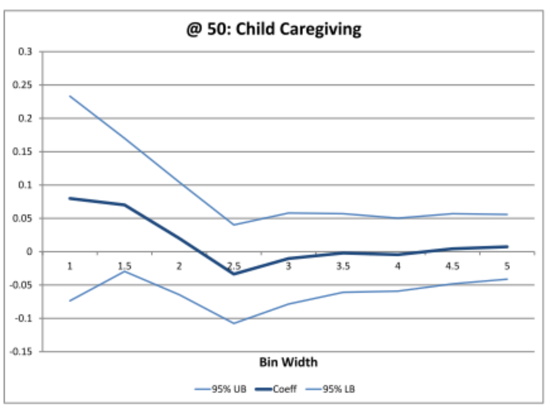
(a) LTC Expenditures



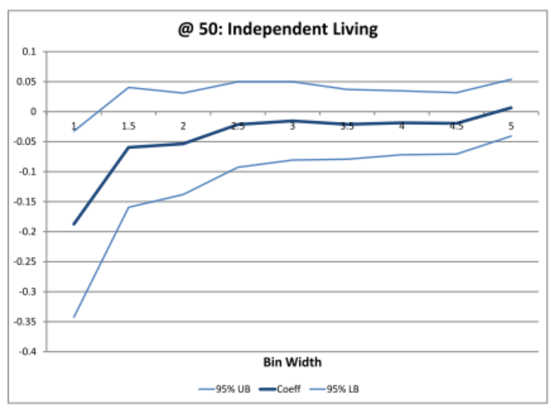
(b) Medical Expenditures



(c) Child Caregiving



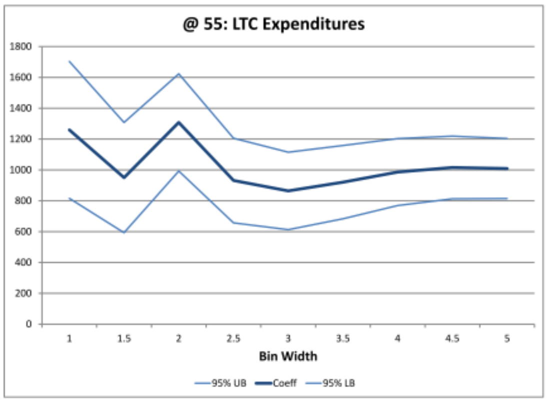
(d) Independent Living



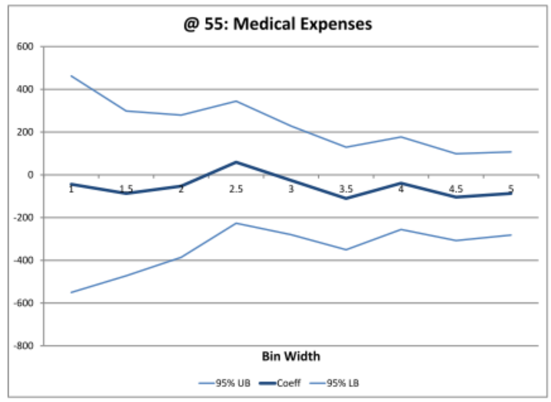
Notes: Each figure displays estimates of β from local linear regression of Equation (1.1) at different bandwidths in increments of 0.5, from 1 to 5. Each figure corresponds to a different outcome.

Figure B.4: Sensitivity to Bandwidth at 55

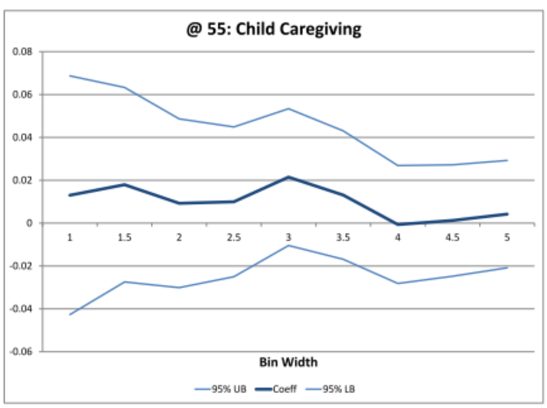
(a) LTC Expenditures



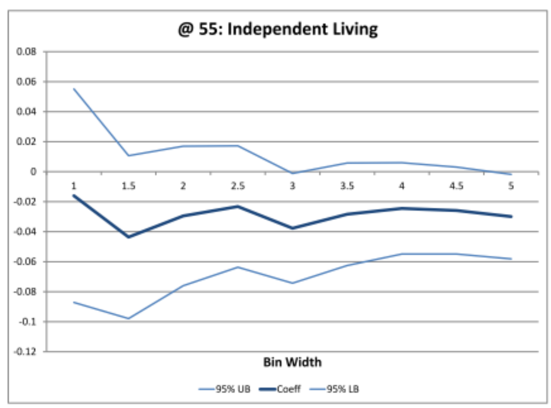
(b) Medical Expenditures



(c) Child Caregiving



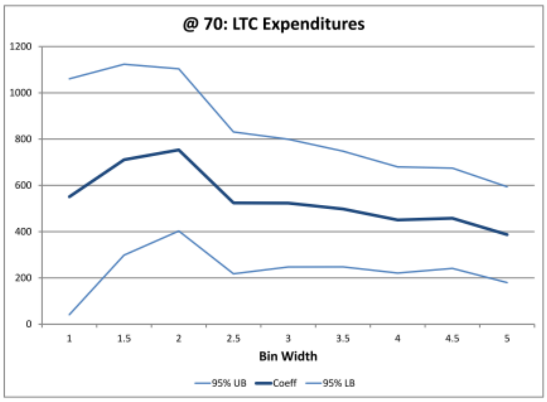
(d) Independent Living



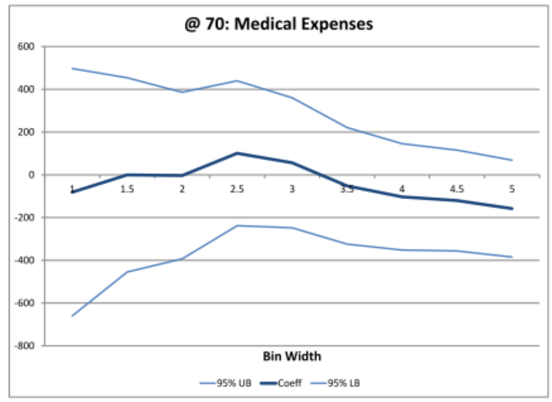
Notes: Each figure displays estimates of β from local linear regression of Equation (1.1) at different bandwidths in increments of 0.5, from 1 to 5. Each figure corresponds to a different outcome.

Figure B.5: Sensitivity to Bandwidth at 70

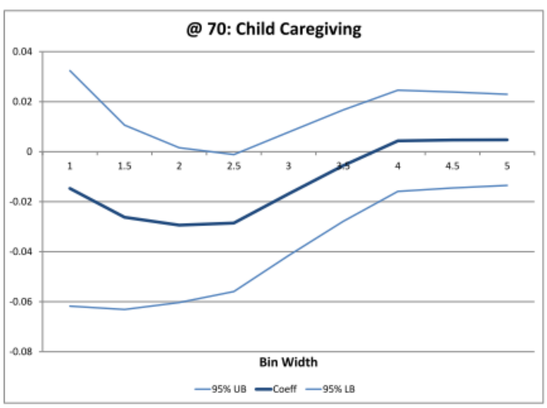
(a) LTC Expenditures



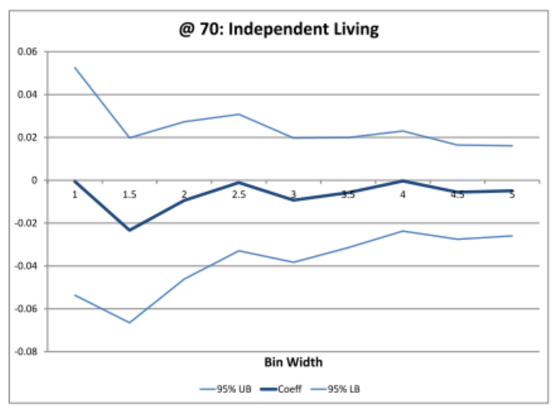
(b) Medical Expenditures



(c) Child Caregiving



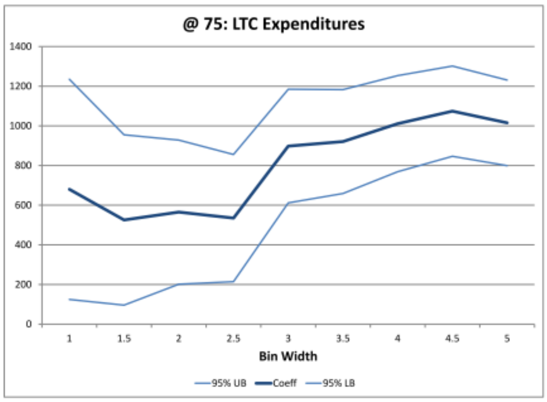
(d) Independent Living



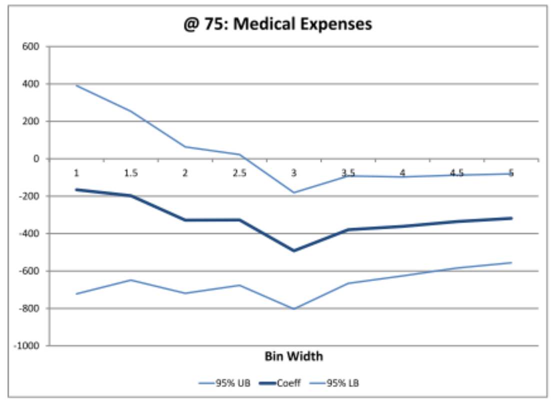
Notes: Each figure displays estimates of β from local linear regression of Equation (1.1) at different bandwidths in increments of 0.5, from 1 to 5. Each figure corresponds to a different outcome.

Figure B.6: Sensitivity to Bandwidth at 75

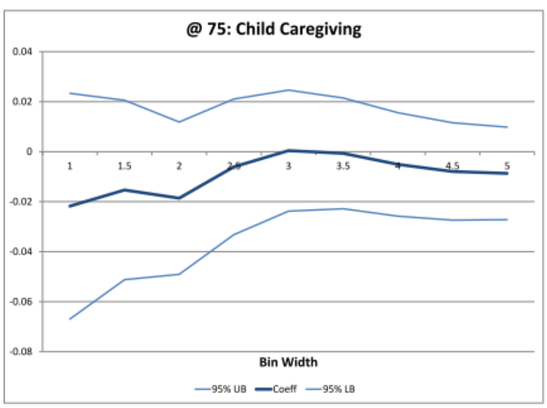
(a) LTC Expenditures



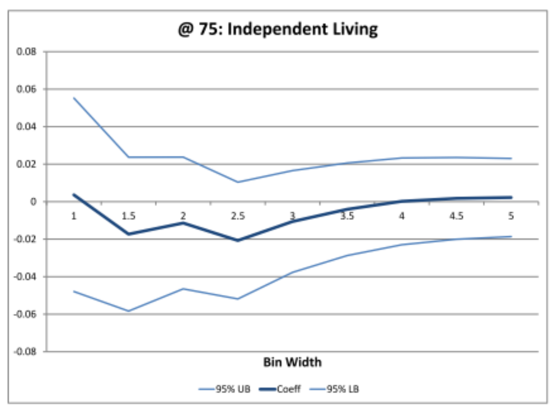
(b) Medical Expenditures



(c) Child Caregiving



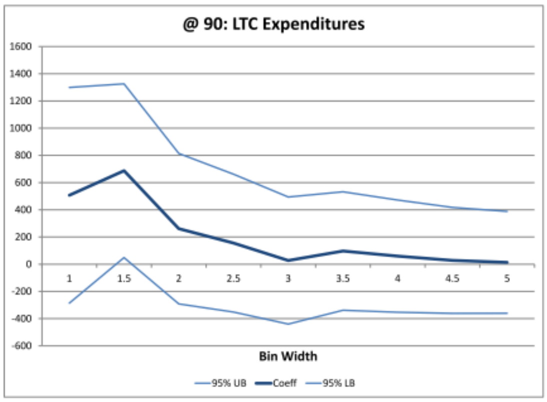
(d) Independent Living



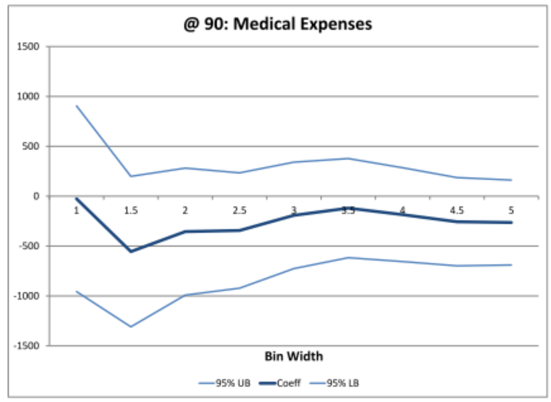
Notes: Each figure displays estimates of β from local linear regression of Equation (1.1) at different bandwidths in increments of 0.5, from 1 to 5. Each figure corresponds to a different outcome.

Figure B.7: Sensitivity to Bandwidth at 90

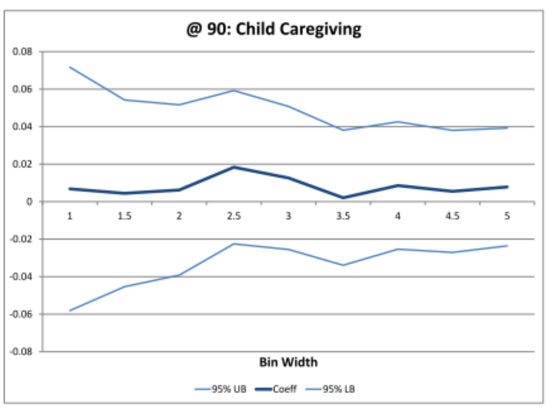
(a) LTC Expenditures



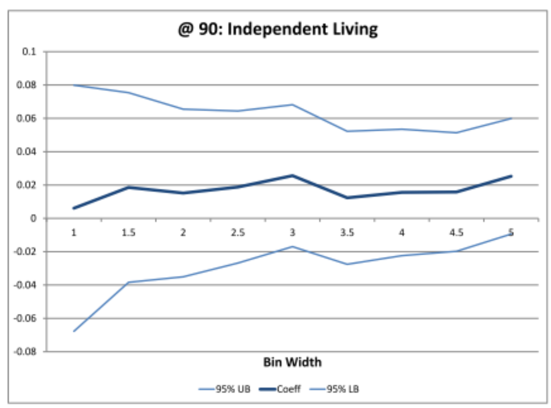
(b) Medical Expenditures



(c) Child Caregiving



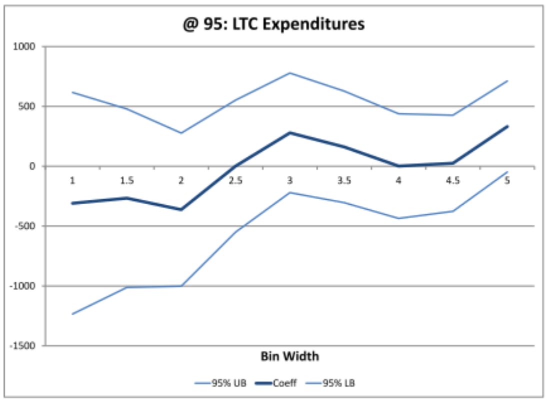
(d) Independent Living



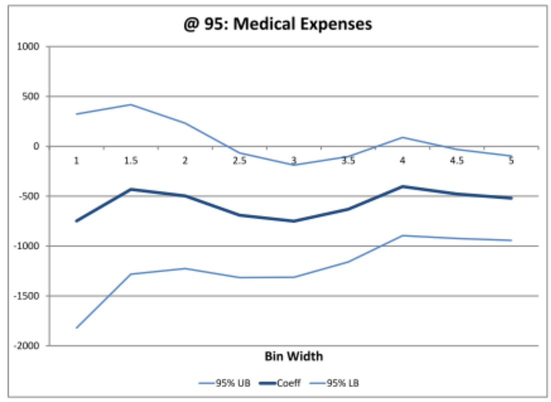
Notes: Each figure displays estimates of β from local linear regression of Equation (1.1) at different bandwidths in increments of 0.5, from 1 to 5. Each figure corresponds to a different outcome.

Figure B.8: Sensitivity to Bandwidth at 95

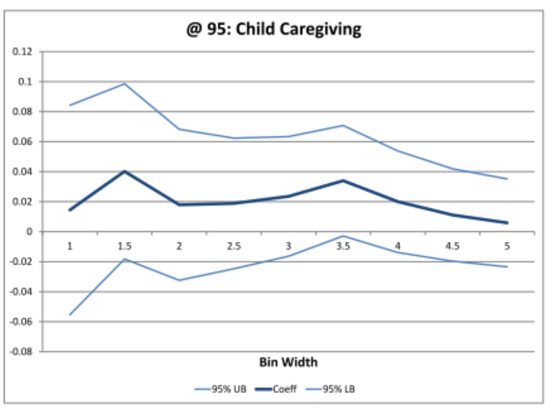
(a) LTC Expenditures



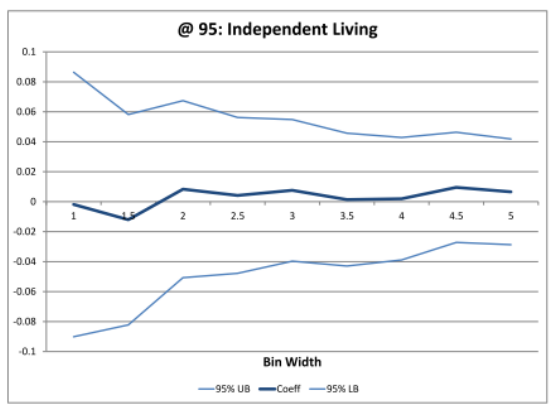
(b) Medical Expenditures



(c) Child Caregiving



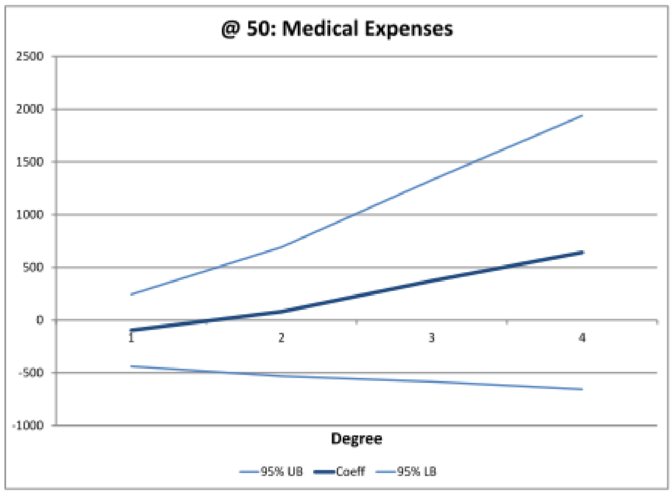
(d) Independent Living



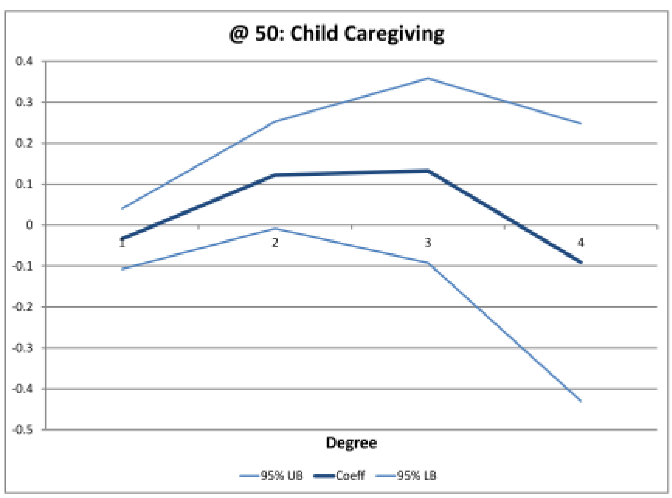
Notes: Each figure displays estimates of β from local linear regression of Equation (1.1) at different bandwidths in increments of 0.5, from 1 to 5. Each figure corresponds to a different outcome.

Figure B.9: Sensitivity to Polynomial Degree at 50

(a) Medical Expenditures



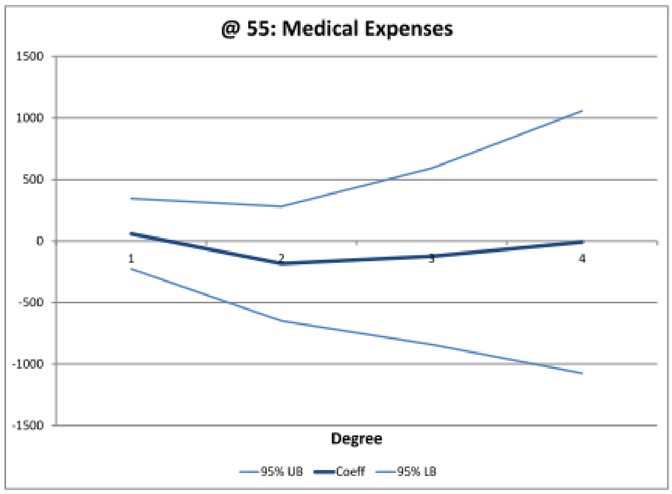
(b) Child Caregiving



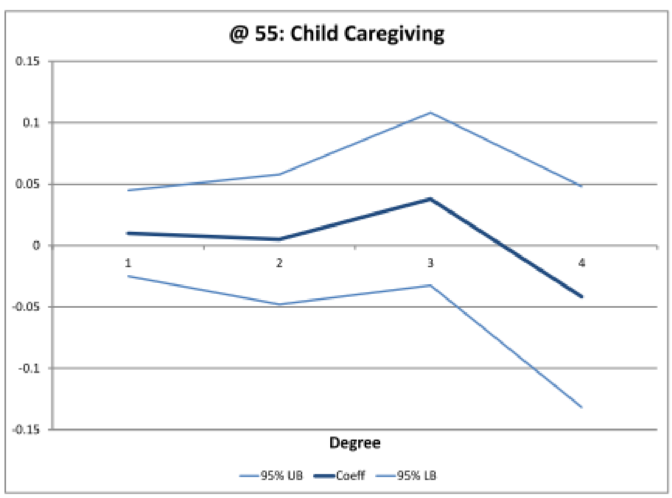
Notes: Each figure displays estimates of β from local linear regression of Equation (1.1) at different polynomial degrees, from 1 to 4. Bandwidth is 2.5. Each figure corresponds to a different outcome.

Figure B.10: Sensitivity to Polynomial Degree at 55

(a) Medical Expenditures



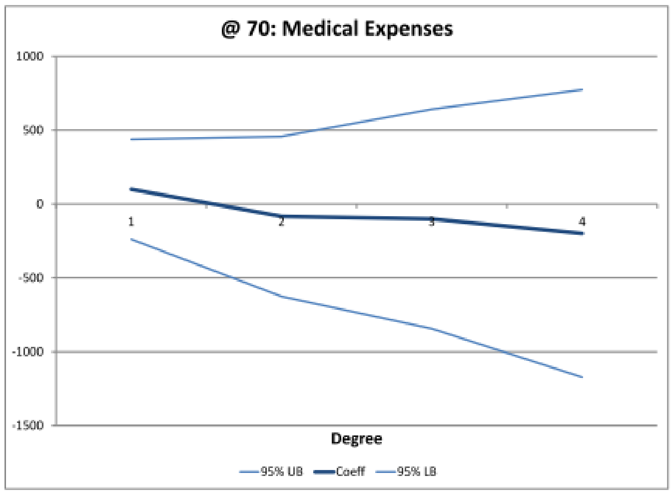
(b) Child Caregiving



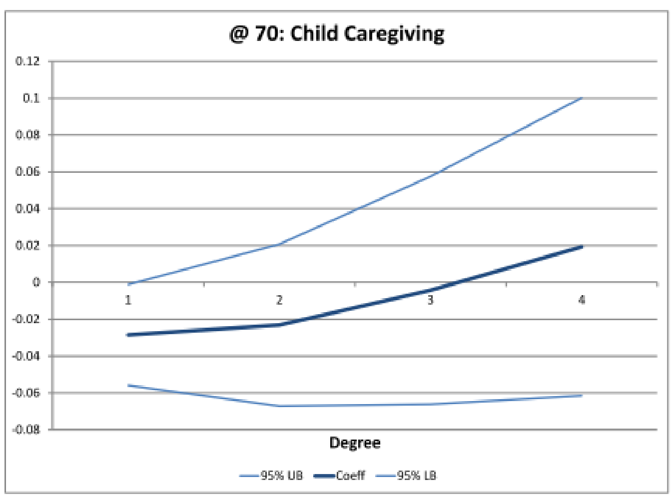
Notes: Each figure displays estimates of β from local linear regression of Equation (1.1) at different polynomial degrees, from 1 to 4. Bandwidth is 2.5. Each figure corresponds to a different outcome.

Figure B.11: Sensitivity to Polynomial Degree at 70

(a) Medical Expenditures



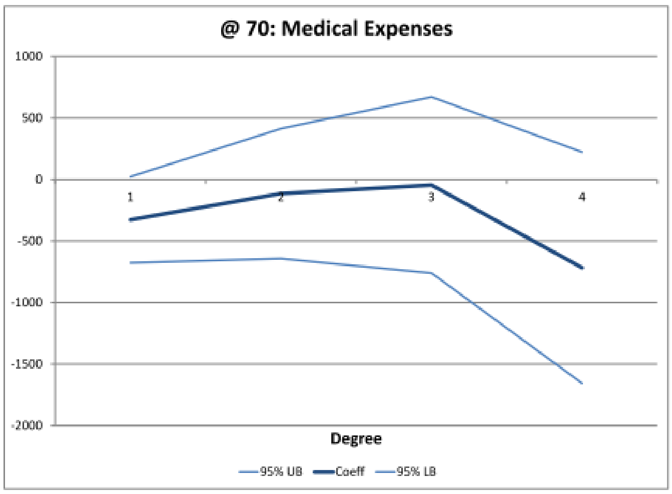
(b) Child Caregiving



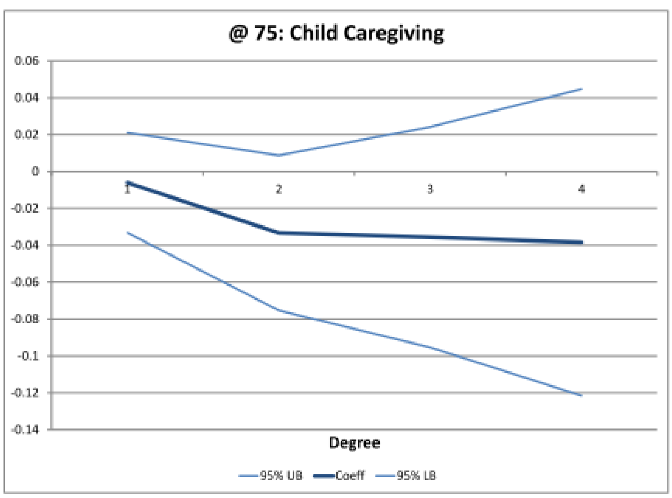
Notes: Each figure displays estimates of β from local linear regression of Equation (1.1) at different polynomial degrees, from 1 to 4. Bandwidth is 2.5. Each figure corresponds to a different outcome.

Figure B.12: Sensitivity to Polynomial Degree at 75

(a) Medical Expenditures



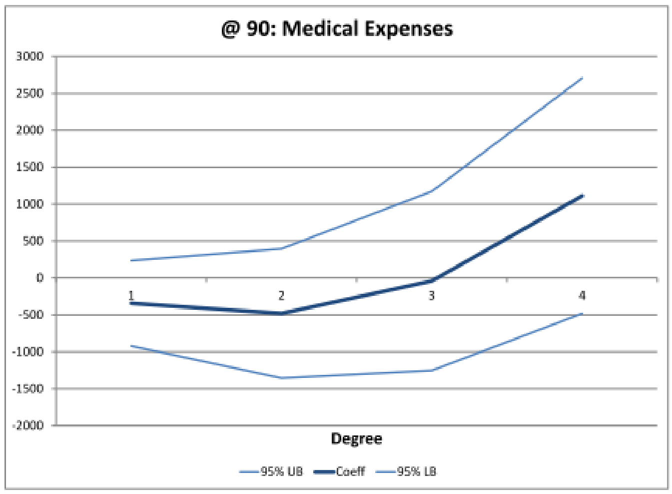
(b) Child Caregiving



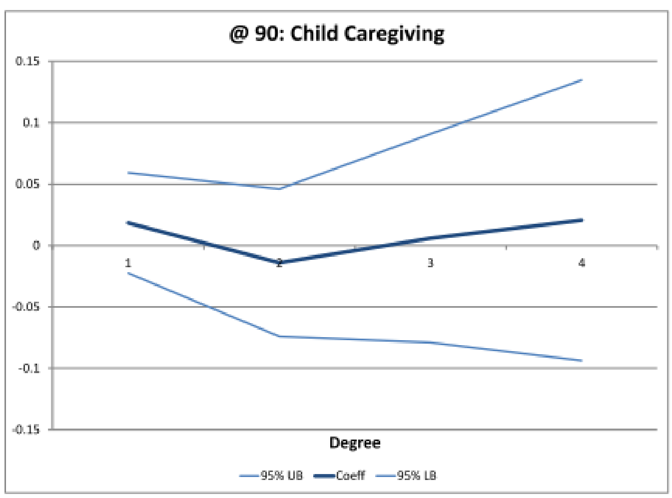
Notes: Each figure displays estimates of β from local linear regression of Equation (1.1) at different polynomial degrees, from 1 to 4. Bandwidth is 2.5. Each figure corresponds to a different outcome.

Figure B.13: Sensitivity to Polynomial Degree at 90

(a) Medical Expenditures



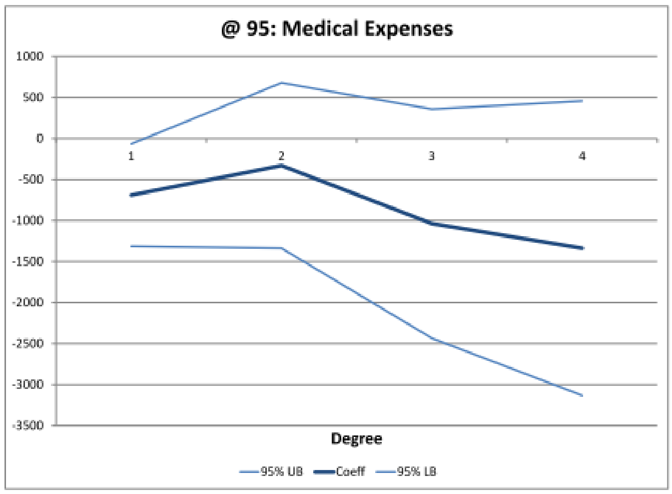
(b) Child Caregiving



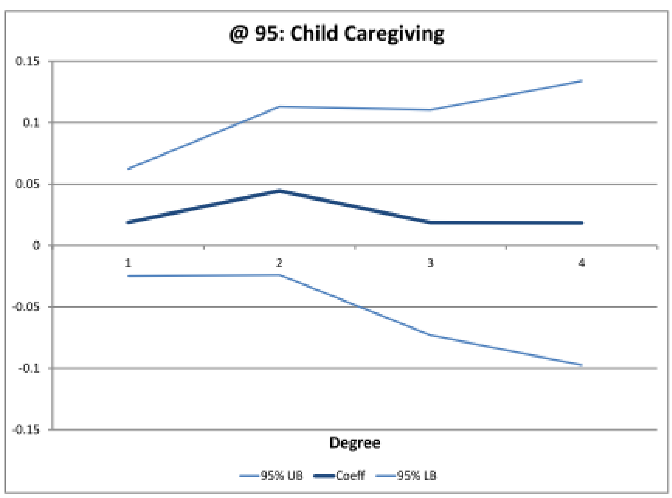
Notes: Each figure displays estimates of β from local linear regression of Equation (1.1) at different polynomial degrees, from 1 to 4. Bandwidth is 2.5. Each figure corresponds to a different outcome.

Figure B.14: Sensitivity to Polynomial Degree at 95

(a) Medical Expenditures



(b) Child Caregiving



Notes: Each figure displays estimates of β from local linear regression of Equation (1.1) at different polynomial degrees, from 1 to 4. Bandwidth is 2.5. Each figure corresponds to a different outcome.