

© 2022 Hyeran Chung

THREE ESSAYS ON HEALTH CARE DISPARITIES IN THE U.S.

BY

HYERAN CHUNG

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Agricultural and Applied Economics
in the Graduate College of the
University of Illinois Urbana-Champaign, 2022

Urbana, Illinois

Doctoral Committee:

Associate Professor Mary Arends-Kuenning, Chair
Associate Professor Flávia Andrade
Associate Professor Yilan Xu
Associate Professor Hope Michelson

ABSTRACT

My dissertation evaluates health disparities in the United States, with specific interests in differences in race, ethnicity, and metro-nonmetro statuses to understand diabetes transitions and management, and employment effects of the influx of foreign-educated registered nurses on native-educated nurses in the U.S. nursing market.

The first chapter of my dissertation examines whether there are racial/ethnic and rurality differences that are associated with effective diabetes management in middle-aged and older U.S. adults with diabetes. My study is among the first in the literature to explore the role of the degree of rurality on diabetes management measured by an objective biomarker, A1c. Using a nationally representative sample of middle-aged and old-aged U.S. adults with diabetes in a longitudinal design, I assess whether there are race/ethnicity and rurality differences that are associated with effective diabetes management and to what extent individual- and county-level determinants explain these differences. My findings are consistent with previous studies, indicating that non-Hispanic Black Americans with diabetes experience ineffective diabetes management compared to non-Hispanic White Americans with diabetes, and the average resident of a nonmetro county has greater odds of showing ineffective diabetes management than the average resident of metro counties. I also find evidence that as the time since the diagnosis increases the quality of health care matters less among individuals with known diabetes. In addition, non-Hispanic Black Americans with known diabetes who live in metro counties are more likely to manage their diabetes than non-Hispanic Black Americans with known diabetes who live in nonmetro counties. The average non-Black female with previously diagnosed diabetes has greater odds of showing effective diabetes management than the average Black female with previously

diagnosed diabetes. Dichotomous threshold measures may not accurately reflect the impact of race, ethnicity, and metro-nonmetro statuses on diabetes management. However, I find little evidence that there are effects of the explanatory variables of interest on the reduction in A1c values between rounds.

The second chapter (co-authored with Dr. Mary Arends-Kuenning and published in the *Journal of Rural Social Sciences* in 2022) explores differences in transitions between stages of type 2 diabetes across racial, ethnic, and urban/rural statuses. Using individual-level factors including biomarkers and self-reported doctor's diagnoses and county-level proxies for structural disadvantage in hazard models, the paper's empirical analysis produces two main findings. First, the likelihood of reporting the transition from prediabetes to diabetes increases with the degree of rurality. Second, race/ethnicity and rurality differences in transitions into diabetes are partially explained by differences in individual-level correlates and in disadvantaged conditions associated with rural counties. Taken together, the two chapters on the topic suggest that future public health efforts designed to reduce racial, ethnic, and metro-nonmetro differences in transitions into diabetes and effective diabetes management should examine the different levels of rurality while controlling for individual-level characteristics and county-level contextual correlates to facilitate regional responses and better resource targeting.

The third chapter (co-authored with Dr. Mary Arends-Kuenning and published in the *IZA Journal of Labor Policy* in 2020) investigates nurse immigration and its impacts on the U.S. nursing labor market with the double instrumentation procedure. Using the 1980–2000 U.S. Censuses and the American Community Survey (ACS) multiyear aggregates, this paper quantifies the displacement effects. Contrary to prior research on migration that typically investigated its effects on the wages or employment of immigrant workers with a single

instrument, we implement a multiple instrumentation procedure to avoid conflating the short- and long-term reactions to the entry of newly arrived foreign-educated nurses. With the double instruments, we find that there is no significant effect of foreign-educated nurses on the employment of native nurses in both the short- and the long-runs. We find marginally significant displacement effects on the youngest native nurses in the long run, whereas highly educated natives are positively affected by the foreign-educated nurses in the short run. We fail to find a negative impact of foreign-educated nurses on native wages but observe that the double IV estimates are larger in magnitude than the corresponding single IV estimates, which implies that researchers should control for dynamic adjustments to past local shocks. Our findings suggest that relying on foreign-educated nurses to fill gaps in the U.S. health care workforce does not harm the employment of native nurses, but rather the entry of foreign-educated nurses to the U.S. nursing market causes an increase in native wages in the long run.

ACKNOWLEDGMENTS

I want to express my gratitude for all the individuals who made this dissertation possible. I am deeply indebted to Mary Arends-Kuenning, who is the ACE advisor I could ever have. I cannot thank her enough for her support, encouragement, and patience. She helped me improve my drafts, clarify ideas, and hone my writing skills. I would also like to extend my deepest gratitude to my committee, Flávia Andrade, Yilan Xu, and Hope Michelson. I am incredibly grateful for their advice and guidance throughout my doctoral career. I must thank Paul Stoddard who taught me how a teacher can motivate students to build up the creativity and model regular interactions with students both in classroom situations and outside the classroom.

I am grateful for the generous support of the Henry Fellowship, Young Fellowship, and Dunn and Linse Fellowship, without which much of this research would not be possible. Thanks also goes out to research funding by the USDA National Institute of Food and Agriculture, Hatch project. I very much appreciate Pamela Splittstoesser, Donna Stites, Melissa Warmbier, and Dung Quach Wisdom for helping me navigate through various administrative processes.

Finally, I would like to thank my family in Korea and the United States who have made my Ph.D. journey possible. To my parents, Moonsuk Chung and Heejoo Lee, I am forever indebted for their love, support, and profound belief in my abilities. My thanks also go to my siblings, Miran Chung, Sungbin Chung, and Sarang Chung. The support I got from these three was unparalleled. I am grateful to my American parents, Linda Walden and Chenson Chen, for always being a stronghold and cheering me up throughout my doctoral studies.

I dedicate this thesis to my best sister, Sarang Chung, who provides me unconditional love and emotional support throughout my doctoral studies.

TABLE OF CONTENTS

CHAPTER 1: EFFECTIVE DIABETES MANAGEMENT AFTER DIAGNOSIS: EXAMINING THE ROLES OF RACE, ETHNICITY, AND METRO-NONMETRO STATUSES.....	1
CHAPTER 2: RACIAL, ETHNIC, AND URBAN/RURAL DIFFERENCES IN TRANSITIONS INTO DIABETES: EVIDENCE FROM THE HEALTH AND RETIRMENT SURVEY BIOMARKER AND SELF-REPORTED DATA.....	54
CHAPTER 3: DO FOREIGN-EDUCATED NURSES DISPLACE NATIVE-EDUCATED NURSES?	85
APPENDIX A: SUPPLEMENTARY MATERIAL FOR CHAPTER 1.....	127
APPENDIX B: SUPPLEMENTARY MATERIAL FOR CHAPTER 3.....	133

CHAPTER 1

EFFECTIVE DIABETES MANAGEMENT AFTER DIAGNOSIS: EXAMINING THE ROLES OF RACE, ETHNICITY, AND METRO-NONMETRO STATUSES

1.1. Introduction

Significant diabetes disparities exist among racial/ethnic and metro-nonmetro statuses in both prevalence and management. With the aging of the U.S. population, these disparities have important public health implications. In 2018 in the U.S., there were 1.5 million new cases of diabetes among adults, or 6.9 per 1,000 persons (Levy, 2021), with the estimated total economic cost of diagnosed diabetes at \$327 billion in 2017 (American Diabetes Association, 2018). This dollar amount included both direct and indirect costs, and it represented a 26% increase for total diabetes-related costs from 2012 cost estimates (American Diabetes Association, 2018). The effects of diabetes on economic burden are mostly due to working-age adults. However, previous studies have shown that the prevalence of diabetes is strongly associated with increased age (Kirkman et al., 2012; Aspray and Unwin, 2006; Durso, 2006; Meneilly and Tessier, 2001). Brown et al. (2003) found that older adults aged 65 and over with diabetes are more likely to face higher mortality, reduced functional status, and increased risk of institutionalization than younger adults. Thus, the focus of this study is U.S. adults over 50.

The burden of diabetes is not shared equally across race and ethnicity. Racial and ethnic minority populations disproportionately exhibit higher rates of diabetes. Resnick et al. (2008) found that the prevalence of diabetes among individuals aged 65 and over is 22.5% in Whites and 35.6% in non-Whites. Unhealthy lifestyle behaviors such as insufficient physical activities (Zhao et al., 2011) and poor nutrition (Aspray and Unwin, 2006) are also risk factors that drive

higher diabetes rates. In addition, geographic differences in diabetes prevalence (Krishna et al., 2010; O'Connor and Wellenius, 2012; Lutfiyya et al., 2009; Lutfiyya et al., 2011) are getting attention by researchers, but empirical explanations for these differences are limited in prior studies. Compared to urban residents, rural residents face health delivery challenges, including limited health care service provision (Panelli et al., 2006; McConnel and Zetzman, 1993) and lack of physicians (Pathman et al., 2006; Gong et al., 2019). Individual-level residency and county-level characteristics themselves related to rurality imply that rural residents are less likely to have regular check-ups (Basu et al., 2004) and to have access to a high quality of health care for diabetes (Chung and Arends-Kuenning, 2022).

The population of U.S. adults over 65 with diabetes is expected to increase by 4.5-fold between 2005 and 2050 (Narayan et al., 2006), which highlights the importance of optimizing diabetes management in older individuals. However, empirical research on effective diabetes management is extremely limited. The present study fills this research gap. Using a nationally representative sample of middle-aged and old-aged U.S. adults with diabetes in a longitudinal design, I assess whether there are race/ethnicity and rurality differences that are associated with effective diabetes management and to what extent individual- and county-level determinants explain these differences. My findings are consistent with previous studies, indicating that non-Hispanic Black Americans with diabetes experience ineffective diabetes management compared to non-Hispanic White Americans with diabetes, and the average resident of a nonmetro county has greater odds of showing ineffective diabetes management than the average resident of metro counties. I also find evidence that as the time since the diagnosis increases the quality of health care matters less among people with known diabetes. In addition, non-Hispanic Black Americans with known diabetes who live in metro counties are more likely to manage their diabetes than

non-Hispanic Black Americans with known diabetes who live in nonmetro counties. The average non-Black female with previously diagnosed diabetes has greater odds of showing effective diabetes management than the average Black male with previously diagnosed diabetes.

My results advance prior studies in several ways. First, I improve upon cross-sectional studies. Using longitudinal data from the Health and Retirement Study allowed me to examine the disparities in diabetes management over time. Second, I take advantage of the rich data available in the Health and Retirement Study to control for both individual- and county-level correlates. Adding individual-level correlates and county-level health care infrastructure indicators may attenuate race, ethnicity, and rurality disparities in diabetes management. Third, my outcome variable is based on an objective measure - derived biomarker data instead of self-reported data that may have the possibility of recall bias. Lastly, I explore those differences in diabetes management among US middle-aged and older-aged adults with diabetes living in three distinct types of counties. My study is among the first in the literature to explore the role of the degree of rurality on diabetes management measured by an objective biomarker, A1c (Mercado et al., 2021).

1.2. Background

1.2.1. Differential effects of metro-nonmetro statuses on diabetes management

Urban-rural disparities have been described for a wide variety of illnesses. Existing research on these differences in health outcomes consistently finds that many rural communities experience greater prevalence of chronic health conditions and complications (Singh et al., 2017; Singh et al., 2015) and less access to comprehensive health care (Hartley, 2004; Lutfiyya et al., 2009; Lutfiyya et al., 2011) than urban communities. However, few studies compare urban-rural

differences in diabetes. Previous studies on urban-rural disparities in U.S. diabetes prevalence found a higher prevalence of diabetes in rural locations when examining binary urban/rural differences (Krishna et al., 2010; O'Connor and Wellenius, 2012). Rural residents with diabetes were more likely to receive inadequate care than non-rural residents (Lutfiyya et al., 2009; Lutfiyya et al., 2011). However, these simplified binary comparisons can mask differences in diabetes across different levels of rurality (James, 2014). In a study of transitions into diabetes using the Health and Retirement Survey biomarker and self-reported data, Chung and Arends-Kuenning (2022) found that individuals in both medium/small metropolitan counties and nonmetropolitan counties were more likely to experience the transition from prediabetes to diabetes than individuals in large metropolitan counties, a distinction that would have been missed had they used a binary metropolitan status measurement.

One of the challenges in understanding metro-nonmetro statuses differences in diabetes management is teasing out whether contextual features of counties account for those differences. Rural communities face inadequate access to primary care and specialty care services compared to urban communities (Pathman et al., 2006). Previous research found that county-level structural disadvantage is an important predictor of health outcomes. For example, lack of access to specialists was the driver of preventable hospitalization rates among rural Medicare beneficiaries with chronic conditions (Johnston et al., 2019). With respect to diabetes development, Chung and Arends-Kuenning (2022) found that an introduction of the index to measure the quality of health care service for diabetes in their models decreased the likelihood of reporting the transition from prediabetes to diabetes.

1.2.2. Racial and ethnic differences in glycemic control

Several recent studies have examined glycemic control outcomes by race and ethnicity. Studies presented their outcomes using point estimates of A1c (an indicator of the average glycemic control over the past 2-3 months) or outcomes relative to a threshold. Regardless of the outcome measure used, differences in glycemic control by racial/ethnic group were observed (Brown et al., 2005; Hausmann et al., 2010; McWilliams et al., 2009; Chew et al., 2008; Kirkbride and Wallace, 2009; Silverman et al., 2011; Saydah et al., 2007; Smalls et al., 2020).

These studies used a variety of data collection methods, including administrative data, laboratory data, and randomized controlled trial from a behavioral intervention. Chew et al. (2008), for example, observed people with diabetes who received continuity of care for at least two years with two outpatient interactions at one of six U.S. public hospitals. Data come from the Consortium hospital systems that were comprised of Grady Health System in Atlanta, Harborview Medical Center in Seattle, Cambridge Health Alliance in Boston, Cook County Bureau of Health Services in Chicago, Louisiana State University Medical Center in Baton Rouge, and the Community Health Network in San Francisco. Chew and her coauthors found African Americans and Hispanic White Americans were less likely than non-Hispanic Whites to have good control ($< A1c$ of 7%). Similarly, Silverman et al. (2011) collected data from the emergency department of a university-affiliated teaching hospital from April 2007 to June 2007. An A1c level was obtained for 880 people with a known history of diagnosed diabetes who presented to the emergency department for acute medical problems. They were consecutively screened for a total of 58 24-hour periods during a 10-week periods with a blood test. Researchers stratified participants' A1c into levels of control, with A1c level greater than or equal to 7% indicating poor control. Silverman et al. (2011) found that minority populations with

a known history of diabetes presenting at a single emergency department are at higher risk for poor glycemic control than non-Hispanic White Americans with previously diagnosed diabetes. Hausmann et al. (2010), using A1c values as their dependent variable with a sample consisting of patients who self-selected into a study of a behavioral intervention, reported that African Americans have higher A1c levels even after controlling for demographic, clinical, and psychosocial factors.

The above studies have limitations. First, Chew et al. (2008) used administrative and patient laboratory data for which the completeness and accuracy are uncertain. For example, not all participating hospital systems were able to provide consistent information about race. Second, the selection of patients from a single emergency department in Silverman et al. (2011)'s study limits generalizability. Third, the study sample may not be representative of the general population of individuals with diabetes in terms of motivation to engage in diabetes self-care behaviors (Hausmann et al., 2010) or utilization of outpatient care (Silverman et al., 2011).

Other studies examining the effects of race and ethnicity on diabetes management use the National Health and Nutrition Examination Survey (NHANES) for a nationally representative sample of the United States. Saydah et al. (2007) and Smalls et al. (2020) analyzed cross-sectional data from NHANES and found non-Hispanic Blacks and Mexican Americans with a previous diagnosis of diabetes were less likely to achieve good glycemic control (A1c levels < 7%) than non-Hispanic Whites. Their models controlled for individual-level correlates that may confound an observed association between race/ethnicity and glycemic control. Although authors found that some of individual-level factors were associated with glycemic control, they could not establish causality because it was a cross-sectional study.

1.3. Data, measurements, and sample statistics

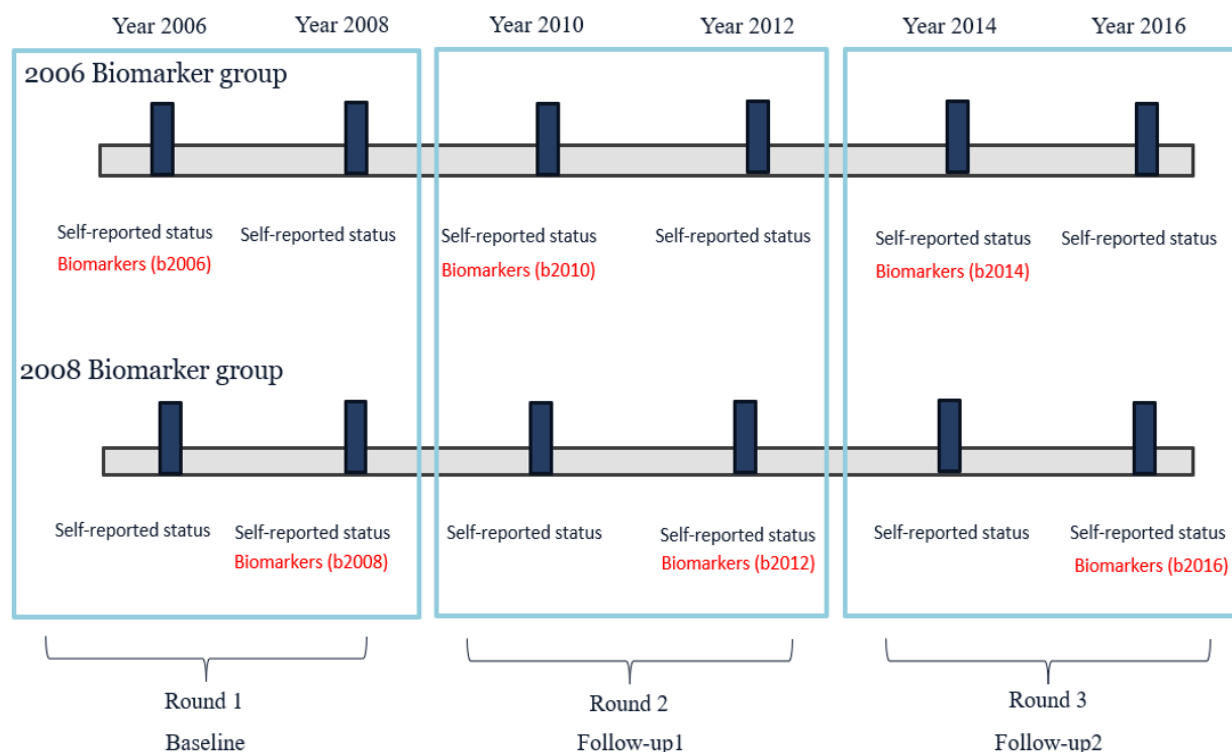
1.3.1. Data

Data for this study come from the 2006-2016 Health and Retirement Study (HRS). The Health and Retirement Study is a biennial longitudinal survey of U.S. households initiated in 1992. It is sponsored by the National Institute on Aging and conducted by the University of Michigan's Institute for Social Research (Juster and Suzman, 1995). The HRS started with a sample of households in which the household heads were 51 to 61 years old and is considered as a representative sample of the entire U.S. population over 50 years of age. As of 2006, the HRS contained 18,469 individuals from AHEAD (born before 1924), CODA (born in 1924-1930), HRS (born in 1931-1941), War Babies (born in 1942-1947), Early Baby Boomers (born in 1948-1951), and Late Baby Boomers (1954-1959). The HRS collects information through face-to-face interviews, phone interviews, and proxy interviews.

Respondents are surveyed for self-reported outcomes biennially, and in addition the HRS expanded to include an enhanced face-to-face (EFTF) interview with a medical examination by highly trained interviewers employed by the Survey Research Center at the Institute for Social Research at the University of Michigan. The medical examination involved measurement of height, weight, mobility, strength, blood pressure, and lung capacity, as well as a collection of biomarker data. Starting in 2006, the HRS randomly selected rotating halves of the panel in each wave for the EFTF interview. The households who were not assigned to an EFTF interview in the 2006 wave were assigned to an EFTF interview in the 2008 wave. In other words, participants receive EFTF interviews every 4 years beginning in 2006/2008. The sample was selected at the household level, and each selected household was required to include at least one age-eligible member. To examine the differences between the samples of those who decided to

provide biomarkers and those who decided against it, I took the baseline data and compared two samples. Non-Hispanic White Americans were more likely to complete the interview and provide the biological samples than non-Hispanic Black Americans or Hispanics. In addition, an individual who is married and highly educated is more likely to provide biological samples is than an individual who is unmarried and with less education than completed high school (Table A.1). Interviewers measured and collected two types of biological samples to evaluate biomarkers, blood and saliva. Saliva is for DNA extraction, and blood is used to measure 5 biomarkers, which are total cholesterol, High-Density-Lipoprotein (HDL) cholesterol, Glycated hemoglobin (A1c), C-reactive protein, and Cystatin C (Crimmins et al., 2013). Interviewees were notified about their biomarker results by mail.

Figure 1.1: Timing of biomarker measurements in the HRS



Source: Chung and Arends-Kuenning (2022)

Figure 1.1 describes the timing of data collection in the HRS panel in two timelines, one for each of the two rotating biomarker groups in the HRS. Self-reported outcomes are recorded every two years, whereas the objective biomarker data is measured every other wave. Each blue square illustrates the baseline, first follow-up, and second follow-up, respectively. Weir (2007) describes how the integration of biomarkers into the HRS validates and adds information to self-reported health, improves modeling of pathways to health, and allows participants to ascertain what they had not previously known about their health.

County-level data came from the Health and Retirement Study Contextual Data Resource (HRS-CDR). The HRS-CDR is a restricted data set that enables researchers to examine the effects of place on health and well-being among the HRS respondents (Ailshire et al., 2020). The Dartmouth Atlas of Health Care in the HRS-CDR documents how medical resources are distributed and used in the United States. Data to identify health professional shortage areas were drawn from the County Characteristics Data set, which is produced by the Inter-university Consortium for Political and Social Research (ICPSR, 2008).

1.3.2. Measurements: diabetes, metro-nonmetro statuses, and health care services

The HRS asks respondents to provide a self-report of diabetes. If a respondent is new in the 2006 wave, the respondent is asked “Has a doctor ever told you that you have high blood sugar or diabetes?” If a respondent participated in a prior HRS wave and reported diabetes in her/his previous interview, the interviewer asks her/him whether s/he wishes to dispute the prior report of illness. If not, the respondent with self-reported doctor’s diagnosed diabetes is asked whether s/he currently takes medication for diabetes. If a respondent participated in a prior HRS interview but did not self-report doctor’s diagnosed diabetes, the respondent is asked “Since we

talked to you, has a doctor told you that you have high blood sugar or diabetes?” If the respondent self-reports doctor’s diagnosed diabetes in the current period, s/he is asked a follow-up question about whether any medication for the illness¹ is taken. I considered a person to have diabetes if s/he self-reported that a doctor diagnosed diabetes. The HRS enhanced interview included a blood draw and blood analysis. As an objective measure of diabetes management, I used the A1c level. The A1c value shows how high an individual’s average blood sugar was over the past three months and has been widely used in clinical settings since the 1980s as a test of glycemic control (Gallagher et al., 2009; Buell et al., 2007; and Ginde et al., 2008). One merit of the A1c test is that the result is insensitive to the timing of measurement and to an individual’s food consumption that day (Chatterji et al., 2012)². Consistent with the American Diabetes Association’s (2014) guidelines, I considered an A1c of lower than 6.5% to be an indicator of effective management of diabetes³. If respondents with no self-reported doctor’s diagnosed diabetes have an A1c value greater than or equal to 6.5%, I defined them as individuals with undiagnosed diabetes.

One of the independent variables of interest is the rurality of the county within which the respondent resides. I used the county as the contextual unit of analysis because the county is small enough to reflect local social and economic conditions (McLaughlin et al., 2001) and is also a relevant geographic unit for the delivery of many social services such as public health. I divided respondents among three levels of rurality based on the Rural-Urban Continuum Codes

¹ This question structure has the advantage of reducing the likelihood of false-positive reporting due to ‘no longer having diabetes’ since the data is updated every two years. Even though diabetes is very unlikely to be completely cured, there is a possibility of complete recovery due to proper management. Update of the status every two years reduces the likelihood of this error.

² The American Diabetes Association recommends that physicians use fasting plasma glucose (FPG) levels to diagnose diabetes, but this test is very sensitive to food consumption (Chatterji et al., 2012).

³ Following guidelines from the American Diabetic Association (2014), individuals with an A1c of lower than 5.7% are normoglycemic, which implies no diabetes. An A1c range of 5.7 to 6.4% identifies individuals with high risk for future diabetes, which means prediabetes. A person with diabetes would have an A1c level of 6.5% or higher.

classified by the U.S. Department of Agriculture's Economic Research Service (USDA ERS 2003)⁴. In this study, (1) Large central/fringe areas are counties in metropolitan areas of 1 million population or more (e.g., Los Angeles County, CA; Cook County, IL), and (2) Medium/small areas are counties in metropolitan areas with populations of less than 1 million (e.g., Champaign County, IL). Although the USDA-ERS separates nonmetro counties into six codes, I combined all nonmetropolitan categories together to ensure large enough sample sizes of people with diabetes in the baseline, hereafter referred to as (3) rural areas. I referred to metropolitan counties as urban areas and nonmetropolitan counties as rural areas.

Taking a contextual perspective, a foundational premise is that the structural contexts surrounding people matter. My county-level contextual influence that may account in part for residential differences in diabetes management is the availability of health care services. The health care infrastructure was measured with two indicators: one is a binary indicator for health professional shortage areas, and the other is the index for the quality of diabetes management. A Health Professional Shortage Area (HPSA) indicates a geographic area in which any part of the county was classified as a shortage area of primary medical care in 2007⁵. The primary factor used to determine a HPSA designation is the number of health professionals relative to the population with consideration of high need⁶. Access to primary health care may facilitate preventive health interventions, and counties assigned as a shortage area may offer insufficient diabetes care to people with diabetes. This is based on a formula derived by the U.S. Health

⁴ The official Office of Management and Budget (OMB) metro and nonmetro categories have been subdivided into three metro and six nonmetro categories. The USDA-ERS distinguished metropolitan counties by the population size of their metro area, and nonmetropolitan counties by adjacency of the county to a metro area and the urban population size.

⁵ I would have included indicators that capture the supply of health professionals for diabetes, but these data are not available at the county level.

⁶ See details: <https://data.hrsa.gov/topics/health-workforce/shortage-areas>

Resources & Services Administration and implemented by requiring updates of HPSAs more than three years old.

To measure the quality of health care focusing on diabetes, I closely followed the way Chung and Arends-Kuenning (2022) created their index for health care quality. They referred to two sets of nationally accepted preventive service guidelines from the American Diabetes Association (ADA) and Healthy People 2020 and determined the threshold. The ADA provided the recommended frequency of each service, and Healthy People 2020 offered national goals. They used the Dartmouth Atlas Data set in 2003-2005, which provides the average annual percentage of people with diabetes aged 65-75 who had at least one of A1c test, eye exam, and blood lipids test in each U.S. county. The county is considered to offer a high-quality health care for diabetes if each service variable exceeded the national goal for the services set by Healthy People 2020⁷. The index takes a value between 0 to 3. One point is added to the index if the average annual percentage of Medicare adults with diabetes aged 65-75 receiving each service is above each threshold. Larger values of the index imply higher quality of care for individuals with diabetes.

1.3.3. Descriptive statistics

Of the 29,842 respondents aged over 50 in the United States (except for Alaska and Hawaii) eligible for the 2006-2016 waves of the HRS interview, 29,835 provided their diabetes status. A total of 20,744 respondents participated in biomarker collection and submitted their A1c values at least once from 2006 to 2016. Among them, 5,682 individuals had ever been

⁷ Both ADA and Healthy People 2020 do not restrict their targets to middle-aged or older adults. To set each threshold of three tests, I created the highest 25% values of each A1c test, eye exam, and blood lipids test among Medicare adults with diabetes aged 65-75. The generated thresholds for each test are similar to the national goals, and I decided to use the national guideline to decide the thresholds (Chung and Arends-Kuenning, 2022).

diagnosed with diabetes over the 2006-2016 waves. Because I aimed to investigate whether individuals with diabetes in the baseline are managing their diabetes well enough by looking at the A1c values in the follow-up, I focused on (1) individuals with diabetes in the baseline and (2) people with diabetes who appeared more than once. Among 5,682 individuals who have ever been diagnosed with diabetes over the 2006-2016 waves, a total of 2,760 in the baseline are not observed in the follow-up. It includes 694 individuals who died between the rounds, and 2,066 individuals who do not provide their A1c values in the follow up and therefore, eventually were lost to follow up. The response rate for biomarker collection between waves is approximately 67%, and I lost over one-third at follow-up. These lead me to include 2,922 people with diabetes. The analyses lastly excluded 116 individuals with diabetes for whom data about their race/ethnicity, SES, health behaviors, and diabetes severity are absent⁸.

I identified two groups of 2,806 individuals with diabetes: 1) people with diagnosed diabetes in round 1 and 2) people who had no diabetes diagnosis in round 1, but had received a diabetes diagnosis between round 1 and round 2. My final sample sizes for each of the study cohorts are 2,521 unique people with previously known diabetes⁹ and 285 people with newly diagnosed diabetes. Among 20,744 respondents who participated in biomarker survey and submitted their A1c values at least once over 2006-2016 waves, the proportion of final sample sizes for individuals with diagnosed diabetes is 13.53% (2,806 unique people) and the proportion of final sample sizes for individuals with undiagnosed diabetes is 2.39% (495 unique people).

Table 1.1 presents the unadjusted mean indicating an effective management rate for

⁸ The number of missing covariates in the baseline are listed: race/ethnicity (missing=10), educational attainment (missing=10), insurance status (missing=2), smoking history (missing=9), BMI information (missing=35), and treatment usage (missing=40).

⁹ Among 2,521 individuals with previously diagnosed diabetes, 1,621 people with diabetes appeared twice, and 900 people with diabetes appeared three times. When an individual with diabetes appears more than twice, I used both rounds of information and took the advantage of a longitudinal data setting.

Table 1.1. Descriptive statistics by metro-nonmetro status

N of unique people	Previously diagnosed (N=3,412)			Newly diagnosed (N=285)		
	2,521 individuals			285 individuals		
	Large (N=1,655)	Med/Small (N=1,046)	Rural (N=720)	Large (N=144)	Med/Small (N=73)	Rural (N=68)
Effective Management	41.93%	40.44%	40.83%	62.50%	64.38%	60.29%
Individual characteristics (Baseline)						
Age 51-59	31.60%	25.24%	23.61%	13.89%	8.22%	10.29%
Age 60-64	15.47%	15.87%	15.42%	21.53%	21.92%	11.76%
Age 65-69	17.95%	18.55%	21.11%	13.89%	24.66%	26.47%
Age 70-74	17.16%	17.21%	19.86%	23.61%	19.18%	25.00%
Age 75+	17.82%	23.14%	20.00%	27.08%	26.03%	26.47%
Female	56.37%	54.49%	59.17%	45.14%	46.58%	54.41%
NHW	49.37%	64.91%	77.36%	68.75%	72.60%	82.35%
NHB	26.53%	16.16%	15.56%	17.36%	12.33%	13.24%
Hispanic	19.70%	16.25%	4.58%	11.11%	9.59%	1.47%
NHO	4.41%	2.68%	2.50%	2.78%	5.48%	2.94%
Educational attainment						
Less than HS	23.32%	30.50%	29.72%	17.36%	21.92%	23.53%
High School (HS)	32.51%	33.08%	37.92%	31.25%	34.25%	44.12%
Above HS	44.17%	36.42%	32.36%	51.39%	43.84%	32.35%
Married	63.50%	67.30%	68.06%	67.36%	69.86%	67.65%
Has Medicare	55.77%	63.10%	67.22%	63.19%	64.38%	80.88%
Visited a doctor	95.50%	95.79%	98.75%	97.22%	95.89%	95.59%
Smoking history						
Never smoked	43.87%	43.79%	46.39%	47.22%	46.58%	48.53%
Former smoker	44.05%	44.74%	41.39%	38.19%	47.95%	38.24%
Current smoker	12.08%	11.47%	12.22%	14.58%	5.48%	13.24%
Overweight	32.75%	31.55%	31.81%	38.89%	32.88%	25.00%
Obese	54.38%	56.31%	58.06%	46.53%	53.42%	60.29%
A1c level	6.93	6.80	6.75	6.34	6.38	6.39
No treatment	19.70%	19.98%	20.00%	36.81%	42.47%	29.41%
Oral only	63.56%	60.61%	60.56%	61.81%	53.42%	66.18%
Insulin only	6.34%	6.88%	8.19%	0.69%	1.37%	1.47%
Both oral and Insulin	10.39%	12.52%	11.25%	0.69%	2.74%	2.94%
County-level characteristics (Baseline)						
Quality of health care (index)	1.89	1.93	1.82	1.88	1.96	1.75
HPSA	47.61%	38.47%	14.64%	45.14%	42.47%	16.18%

Note: Effective Management indicates A1c<6.5% in the follow-up.

Abbreviations: NHW = non-Hispanic Whites; NHB = non-Hispanic Blacks; NHO = non-Hispanic other races; HPSA = Health Professional Shortage Area

diabetes and values of predictor variables across the metro-nonmetro statuses. The average unadjusted effective management rate indicator for diabetes is the lowest for nonmetro (rural counties for people regardless of the timing of diagnosis. Among those with previously diagnosed diabetes in the baseline, individuals living in rural counties were less likely to have higher education and more likely to be married compared to those living in metro counties. With respect to health behaviors, rural residents who were previously diagnosed in the baseline were more likely to be current smokers and to be obese than nonmetro residents with previously diagnosed diabetes were. In terms of the characteristics of counties themselves, urban counties were more likely to be designated as primary care health professional shortage areas than rural counties. The designation might not characterize an actual shortage but might still allow for federal program support for the area (Monnat and Pickett, 2011). To help reverse the pattern that primary care physician's density in urban counties is double that of rural counties (Machado et al., 2021), the federal government provides incentives, such as loan repayment programs, bonus payment programs, or programs to recruit noncitizen foreign medical graduates, for physicians practicing in primary care health professional shortage areas (Scannell et al., 2021). Of the rural physician staff in 2004, 44.5% are enrolled in one of these programs, almost twice the average proportion in urban health centers (Rosenblatt et al., 2006).

In terms of individuals' characteristics in the baseline among people who had no diabetes diagnosis in round 1, but had a diabetes diagnosis in subsequent rounds, over three quarters of rural residents are 65 and over, whereas less than 70% of residents in metro counties are 65 and over. Rural residents with newly diagnosed diabetes are also less likely to manage their weight than their counterparts in metro areas. There is a clear metro-nonmetro gradient in blood sugar, with the overall A1c values increasing across the gradient from the most metro to rural counties.

Whereas over 15% of individuals with previously diagnosed diabetes depended on insulin treatment in the baseline, over 95% of individuals with newly diagnosed diabetes got no treatment or oral treatment only in the baseline. It implies that people with newly diagnosed diabetes are prone to have less severe diabetes than those with previously diagnosed diabetes. In terms of the characteristics of counties, rural counties had the lowest index for the quality of health care and lowest percentages of health professional shortage areas compared to other urban and medium-size counties.

Table 1.2 presents the unadjusted mean indicating the effective management rate for diabetes for each racial/ethnic group by gender and rurality. Non-Hispanic White women with diabetes are more likely to manage their diabetes than non-Hispanic White men with diabetes regardless of the degree of rurality. Among those with previously diagnosed diabetes in medium/small metro counties, non-Hispanic Black men were able to manage their diabetes more effectively than non-Hispanic Black women. For Hispanic Americans with previously diagnosed diabetes, women living in large metro counties were more likely to manage their diabetes than men living in large metro counties. Panel B in Table 1.2 shows descriptive statistics for individuals with newly diagnosed diabetes. Some of the gender, race/ethnicity, and residence groups are very small. Non-Hispanic Black women and Hispanic women living in large metro counties were less likely to manage their diabetes well than non-Hispanic Black men and Hispanic men living in large metro counties, respectively.

Table 1.2. Descriptive statistics for each racial and ethnic group by gender and rurality

Panel A: Previously diagnosed (N=3,412)						
Number of unique people	2,521 individuals					
	Large (N=1,655)		Med/Small (N=1,046)		Rural (N=720)	
	Male (N=722)	Female (N=933)	Male (N=476)	Female (N=570)	Male (N=294)	Female (N=426)
<u>An effective management (A1c<6.5%) by Race/Ethnicity</u>						
NHW	43.88% (N=392)	53.18% (N=425)	45.86% (N=338)	46.04% (N=341)	37.96% (N=245)	45.83% (N=312)
NHB	26.35% (N=167)	35.29% (N=272)	40.68% (N=59)	30.91% (N=110)	38.24% (N=34)	39.74% (N=78)
Hispanic	37.80% (N=127)	39.20% (N=199)	27.14% (N=70)	25.00% (N=100)	25.00% (N=8)	20.00% (N=25)
NHO	41.67% (N=36)	40.54% (N=37)	11.11% (N=9)	42.11% (N=19)	57.14% (N=7)	27.27% (N=11)
Panel B: Newly diagnosed (N=285)						
Number of unique people	285 individuals					
	Large (N=144)		Med/Small (N=73)		Rural (N=68)	
	Male (N=79)	Female (N=65)	Male (N=39)	Female (N=34)	Male (N=31)	Female (N=37)
<u>An effective management (A1c<6.5%) by Race/Ethnicity</u>						
NHW	59.68% (N=62)	62.16% (N=37)	56.67% (N=30)	65.22% (N=23)	54.17% (N=24)	62.50% (N=32)
NHB	77.78% (N=9)	62.50% (N=16)	66.67% (N=3)	83.33% (N=6)	40.00% (N=5)	100.00% (N=4)
Hispanic	71.43% (N=7)	66.67% (N=9)	40.00% (N=5)	100.00% (N=2)	0.00% (N=0)	100.00% (N=1)
NHO	100.00% (N=1)	33.33% (N=3)	100.00% (N=1)	100.00% (N=3)	50.00% (N=2)	0.00% (N=0)
Abbreviations: NHW = non-Hispanic Whites; NHB = non-Hispanic Blacks; NHO = non-Hispanic other races						

1.4. The empirical model

To examine the differences in diabetes management, I estimated the following equation:

$$Y_{ijt_f} = \beta_0 + \beta_1 Indiv_{ijt_b} + \beta_2 County_{jt_b} + W_{it_f} + \varepsilon_{ijt}$$

where Y is equal to 1 when an individual i with diagnosed diabetes living in j county in the baseline does not exceed 6.5% of the A1c level in the follow-up, t_f . This is defined as effective diabetes management. For the alternative regression to assess effective diabetes management, I additionally defined the dependent variable as the change in A1c levels between rounds, indicating improvement in severity of diabetes. This response variable is continuous and equal to the differences in A1c values of an individual i with diagnosed diabetes living in j county in the baseline and A1c values of an individual i with diagnosed diabetes living in j county in the follow-up, t_f . I focused on 2,806 individuals with diagnosed diabetes in the baseline who responded to the survey more than once. It is because I aimed to investigate the A1c level in the follow-up to estimate whether people with diagnosed diabetes in the baseline are managing their diabetes adequately. I excluded all counties in Alaska and Hawaii due to missing values on county-level predictors. Therefore, analyses included 44 states¹⁰ and 510 of the 3,142 U.S. counties.

The regressions included the individual-level characteristics ($Indiv_{ijt}$) and the county-level contextual variables ($County_{jtb}$) which, if excluded, might have caused bias in the estimated coefficients. Model 1 controlled for the degree of rurality of residence, U.S. Census regions of residence, and demographic factors. The counties of residence are classified into three types to assess the effects of geographic variation on diabetes care. The Beale Codes assigned in the baseline were used to categorize counties using a three-point graduated scale with 1 being the most urban and 3 being the most rural, according to the USDA-ERS (2003)¹¹. The following race/ethnicity groups were included: non-Hispanic Whites (reference group), non-Hispanic

¹⁰ Four states where no individual with diagnosed diabetes were observed in the baseline were excluded: Montana, Rhode Island, South Dakota, and Utah.

¹¹ I examined Urban Influence Codes (UICs) as well as the RUCCs. The classification into the three categories of the rural/urban continuum changed only trivially and using the UICs would result in a loss of observations because county FIPS were missing in the data set.

Blacks, non-Hispanic other races, and Hispanics. Other demographic factors included were gender (male = reference group) and dummy variables for age groups: 51-59 (reference category), 60-64, 65-69, 70-74, and 75 and over.

In model 2, I added several individual-level variables measured in the baseline to examine the extent to which the metro-nonmetro differences are explained by differences in individual-level characteristics and whether the socioeconomic status and health behaviors of individuals explain differences in diabetes maintenance in large central/fringe counties and those in medium/small counties or rural counties. Individual-level SES indicators included dummies for educational attainment, marital status, and Medicare coverage in the baseline, Health behavior indicators included dummies for whether the respondent reported visiting a doctor in the last two years, smoking status, and overweight and obese weight status, all reported in the baseline. Educational attainment was measured using three categories: less than high school graduate, high school graduate (reference group), and above high school degree. Dichotomous variables for marital status (married = reference category), and Medicare coverage (having Medicare coverage = reference category) in the baseline were included as controls. Smoking status was classified into three categories: current smokers, former smokers, and people who never smoked (reference category). Being overweight and being obese were dichotomous variables determined by Body Mass Index (BMI). A BMI from 25 to <30 was classified as overweight, and a BMI of 30 or higher was classified as obese.

Model 3 added the individual's baseline A1c value and self-reported medical treatment to proxy for severity of diabetes in the baseline. Treatment was classified into four categories: no treatment (reference group), oral treatment only, insulin only, and both oral and insulin treatment. The more severe the illness is, the stronger the prescribed treatment is. Lastly, I added

contextual variables that proxy for county-level health care disadvantages in model 4. The health care infrastructure is a potentially important contextual factor. Lack of health care services may result in missed opportunities to detect diabetes and postpone timely medical treatment (Zhang et al., 2008). Rural areas face significant challenges in availability and accessibility to needed care. Rural residents are 10-20% less likely than urban residents to receive preventive screenings (Michelle et al., 2000). In addition, minorities with diabetes were less likely to have a dilated ophthalmological examination and a lipid profile, compared to non-Hispanic White Americans (Thackeray et al., 2004). Limited choices in health services result in underdiagnosis and less optimal health outcomes (Saaddine et al., 2002; Mainous et al., 2004). Thus, I included a binary indicator for whether any county was classified as a health care professional shortage area and an index for the quality of diabetes care, which takes a value between 0 to 3 to adjust for health care infrastructure. I also included the wave (time) fixed effect, W_{itf} , and standard errors are clustered at the individual level in all regressions.

1.5. Results for estimation of odds of having effective diabetes management

1.5.1. Correlates of racial, ethnic, and metro-nonmetro differences with effective diabetes management among individuals with known diabetes

Table 1.3 presents four models estimating odds of having effective management of diabetes among people with previously diagnosed diabetes. The results in Model 1 demonstrate that as the degree of rurality of the county increases, the probability of effective management for their residents with diabetes decreases. Compared to the average resident in large central/fringe counties with a previous diabetes diagnosis, the average resident in medium/small metro counties with diagnosed diabetes has about 17.1% (OR=0.829) lower odds of having effective

management, and the average resident in rural counties with diagnosed diabetes has about 21% (OR= 0.790) lower odds of having effective management. There are statistically significant differences in the race/ethnicity coefficients, suggesting that non-Hispanic Black and Hispanic individuals with diabetes who were previously diagnosed are less likely to have effective diabetes management than non-Hispanic White Americans are.

The results in Model 2 that add individual-level indicators of socioeconomic status (SES) and health behaviors to the baseline in Model 1 show that the addition of these variables to control for individual-level characteristics do not explain the ethnic, racial, and metro-nonmetro differences in effective management of diabetes, as evidenced by statistically significant and non-remarkable change in the magnitudes of their coefficients. The individual-level health conditions that affect diabetes care are being a former smoker and being obese. It is documented that smoking increases risk for incident diabetes mellitus (Nakanishi et al., 2000; Yeh et al., 2010; Oba et al., 2012). In contrast to my expectation that experience of smoking has negative effects on effective diabetes management, I found that the average former smoker with previously diagnosed diabetes was 15.2% more likely to manage his/her diabetes than the average non-smoker with known diagnosed diabetes. It may be related to an individual's will power and commitment to pursue good health after smoking cessation. In addition, Voulgari et al. (2011) found that smoking cessation is associated with a reduction in levels of albuminuria and improvement of glycemic control and lipid profile, indicating that the negative impacts of smoking are reversible. People who were classified as obese in the baseline were 19.2% less likely to manage their diabetes than people who were not obese. Negative effects of being obese on effective diabetes management is marginally significant and becomes statistically insignificant with the introduction of disease severity in the baseline.

Neither does the introduction of individual-level severity of disease result in a dramatic reduction in the estimates of the magnitudes of the county-type variables (Model 3). Compared to a non-Hispanic White American, a Black American has about 26.1% (OR=0.739) lower odds of having effective management, and a Hispanic American has about 20.3% (OR= 0.797) lower odds of having effective management after controlling for baseline severity of diabetes. The results also show that as the degree of severity of diabetes in the baseline increases, the probability of effective management for their diabetes decreases.

Table 1.3. Odds ratio of having effective diabetes management among individuals with previously diagnosed diabetes

	(1)	(2)	(3)	(4)
Dep=1 when an individual with diabetes in the baseline does not exceed 6.5% of the A1C level in the follow-up.				
<i>Rurality (ref=Large central/fringe counties)</i>				
Medium/small counties (Baseline)	0.829** (0.070)	0.835** (0.071)	0.805** (0.074)	0.804** (0.074)
Rural counties (Baseline)	0.790** (0.077)	0.793** (0.078)	0.766** (0.082)	0.768** (0.084)
<i>Race/Ethnicity</i>				
Hispanics	0.590*** (0.065)	0.587*** (0.070)	0.797* (0.105)	0.804* (0.107)
NHB	0.571*** (0.055)	0.554*** (0.055)	0.739*** (0.080)	0.747*** (0.082)
Female (ref=male)	1.217*** (0.088)	1.241*** (0.098)	1.129 (0.097)	1.130 (0.097)
Individual-level SES and health behavior (Baseline)				
<i>Education (ref=HS)</i>				
Less than HS		1.136 (0.111)	1.222* (0.131)	1.223* (0.132)
Above HS		1.177* (0.100)	1.091 (0.101)	1.088 (0.101)
Married		0.897 (0.073)	0.886 (0.079)	0.885 (0.079)
Has Medicare		1.194 (0.143)	1.222 (0.165)	1.222 (0.165)

Table 1.3. Odds ratio of having effective diabetes management among individuals with previously diagnosed diabetes (cont)

Visited a doctor	1.191 (0.256)	1.037 (0.253)	1.039 (0.253)
<i>Smoking status (ref=Never smoked)</i>			
Former smoker	1.152* (0.091)	1.162* (0.100)	1.164* (0.100)
Current smoker	1.168 (0.142)	1.067 (0.144)	1.069 (0.144)
Overweight	1.007 (0.120)	1.106 (0.145)	1.107 (0.146)
Being obese	0.808* (0.093)	1.033 (0.131)	1.034 (0.131)
Severity of diabetes (Baseline)			
A1c level		0.507*** (0.023)	0.507*** (0.023)
<i>Treatment (ref=None)</i>			
Oral only		0.436*** (0.044)	0.436*** (0.044)
Insulin only		0.277*** (0.054)	0.278*** (0.055)
Both oral and Insulin		0.243*** (0.041)	0.243*** (0.041)
County-level contextual factors (Baseline)			
Quality of Health Care			1.060** (0.126)
HPSA			0.971 (0.089)
Wave FE	X	X	X
Numbers of observation	3,421	3,421	3,421

Abbreviations: NHW = non-Hispanic Whites; NHB = non-Hispanic Blacks; NHO = non-Hispanic other races; HPSA = Health Professional Shortage Area

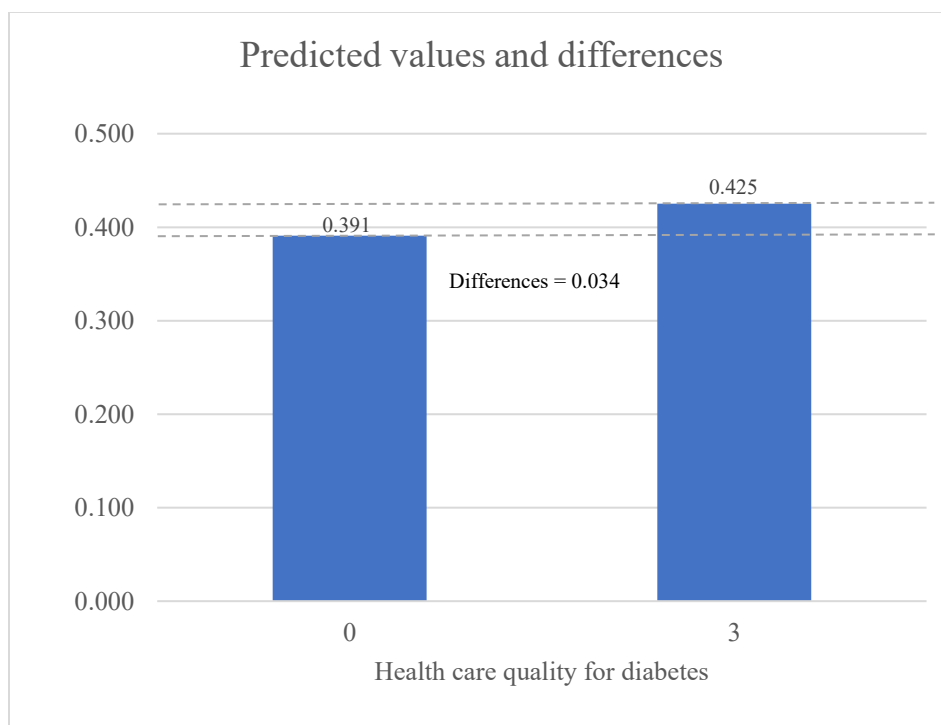
Note: All models include dummies for age groups (51-59, 60-64, 65-69, 70-74, and 75 and over), and the four regions in the U.S. Census, with South as the reference category.

Note: Standard errors in parentheses are clustered at the individual level. Significance is indicated by the symbol * at the 10% level, ** at the 5% level, and *** at the 1% level.

In Model 4, I added the county-level contextual features to the model. The finding suggests a reduced difference in the likelihood of effective management between the average individual in a large central/fringe metro county and the average individual living in a

medium/small metro county or a nonmetro county. In particular, after controlling for county-level characteristics of healthcare services, the average individual living in a medium/small metro county had about 19.6% lower odds of having effective management, and the average individual living in a rural county had about 23.2% lower odds of having effective management compared to the average individual living in a large central/fringe county. Introducing the county-level contextual variables additionally decreases the racial/ethnic differences in effective management for minorities. The average Black American with diabetes who was previously diagnosed has about 25.3% lower odds of having effective management than an average non-Hispanic White American with diabetes. It is marginally significant, but the average Hispanic American with diabetes is less likely to manage his/her diabetes than the average non-Hispanic White American with diabetes.

Figure 1.2. Predicted values and difference in the means to measure the effects of health care quality in main regression analysis



The only county-level contextual variable that approached statistical significance was the quality of health care for diabetes. The result suggests that going from quality 0 to quality 3 increases the probability of having effective diabetes management by 0.034, or 3.41 percentage points (from 0.391 to 0.425), or by 8.02 percent¹² (Figure 1.2). Therefore, racial, ethnic, and geographic differences in diabetes management are partially explained by differences in individual-level determinants and further explained by differences in county-level contextual features.

1.5.2. Correlates of racial, ethnic, and metro-nonmetro differences with effective diabetes management among individuals with known diabetes, stratified by treatment regimens

Models including all the controls discussed above were used to present correlates of differences with effective diabetes management, stratified by medical treatment regimens reported in the baseline (Table 1.4). The degree of rurality was significantly associated with lower likelihood of having effective management among individuals with previously diagnosed diabetes who take no treatment in the baseline. Compared to the average individual with diabetes who had no treatment in the baseline living in a large central/fringe metro county, the average individual with diabetes who had no treatment in the baseline living in a medium/small metro county has about 44.5% lower odds of having effective management (OR=0.555). The average individual with diabetes with no treatment in the baseline living in a rural county is less likely to show effective management (OR=0.579) than the average individual with diabetes who had no treatment in the baseline living in a large central/fringe metro county. I found statistically

¹² Because of the nonlinearity of the logit model, I calculated the predicted values when the healthcare quality is equal to either 0 or 3. I evaluated them at the mean of all the other independent variables and allowed only health care quality to vary. Then, I took the difference in the mean of each predicted value. Both predicted values range between 0 and 1, and the difference in the means is positive (Chung and Arends-Kuenning, 2022).

significant racial/ethnic disparities among individuals with diabetes who had oral treatment only in the baseline. Both the average Hispanic American and non-Hispanic Black American receiving only oral treatment in the baseline have about 28.7% and 26% lower odds of having

Table 1.4. Odds ratio of having effective management for diabetes among people with previously diagnosed diabetes, stratified by treatment regimens in the baseline

	(1) All [N=3,421]	(2) No T [N=679]	(3) Oral T only [N=2,122]	(4) Insulin+ [N=620]
Dep=1 when an individual with diabetes in the baseline does not exceed 6.5% of the A1C level in the follow-up.				
<i>Rurality (ref=Large central/fringe counties)</i>				
Medium/small counties (Baseline)	0.804** (0.074)	0.555*** (0.123)	0.873 (0.098)	0.939 (0.252)
Rural counties (Baseline)	0.768** (0.084)	0.579** (0.149)	0.848 (0.114)	0.698 (0.229)
<i>Race/Ethnicity</i>				
Hispanics	0.804* (0.107)	0.885 (0.317)	0.713** (0.114)	0.978 (0.368)
NHB	0.747*** (0.082)	0.769 (0.218)	0.740** (0.099)	0.642 (0.198)
Female (ref=male)	1.130 (0.097)	0.953 (0.194)	1.172 (0.124)	1.161 (0.302)
Individual-level SES and health behavior (Baseline)				
<i>Education (ref=HS)</i>				
Less than HS	1.223* (0.132)	0.788 (0.204)	1.364* (0.181)	1.369 (0.394)
Above HS	1.088 (0.101)	0.963 (0.214)	1.144 (0.130)	1.186 (0.336)
Married	0.885 (0.079)	0.734 (0.157)	0.877 (0.096)	1.341 (0.362)
Has Medicare	1.222 (0.165)	0.788 (0.257)	1.289 (0.218)	1.907** (0.649)
Visited a doctor	1.039 (0.253)	1.065 (0.589)	0.914 (0.287)	1.959 (1.433)
<i>Smoking status (ref=Never smoked)</i>				
Former smoker	1.164* (0.100)	1.218 (0.256)	1.061 (0.111)	1.596 (0.398)

Table 1.4. Odds ratio of having effective management for diabetes among people with previously diagnosed diabetes, stratified by treatment regimens in the baseline (cont)

Current smoker	1.069 (0.144)	0.901 (0.257)	1.044 (0.180)	1.751 (0.673)
Overweight	1.107 (0.146)	1.159 (0.325)	1.19 (0.203)	0.721 (0.280)
Being obese	1.034 (0.131)	0.931 (0.256)	1.160 (0.190)	0.818 (0.293)
Severity of diabetes (Baseline)				
A1c level	0.507*** (0.023)	0.327*** (0.046)	0.533*** (0.029)	0.580*** (0.056)
County-level contextual factors (Baseline)				
Quality of Health Care	1.060** (0.126)	1.085 (0.305)	1.074* (0.156)	0.786 (0.264)
HPSA	0.971 (0.089)	0.891 (0.194)	1.029 (0.116)	1.166 (0.308)
Wave FE	X	X	X	X

Abbreviations: NHW = non-Hispanic Whites; NHB = non-Hispanic Blacks; NHO = non-Hispanic other races; HPSA = Health Professional Shortage Area

Note: All models include dummies for age groups (51-59, 60-64, 65-69, 70-74, and 75 and over), and the four regions in the U.S. Census, with South as the reference category.

Note: Standard errors in parentheses are clustered at the individual level. Significance is indicated by the symbol * at the 10% level, ** at the 5% level, and *** at the 1% level.

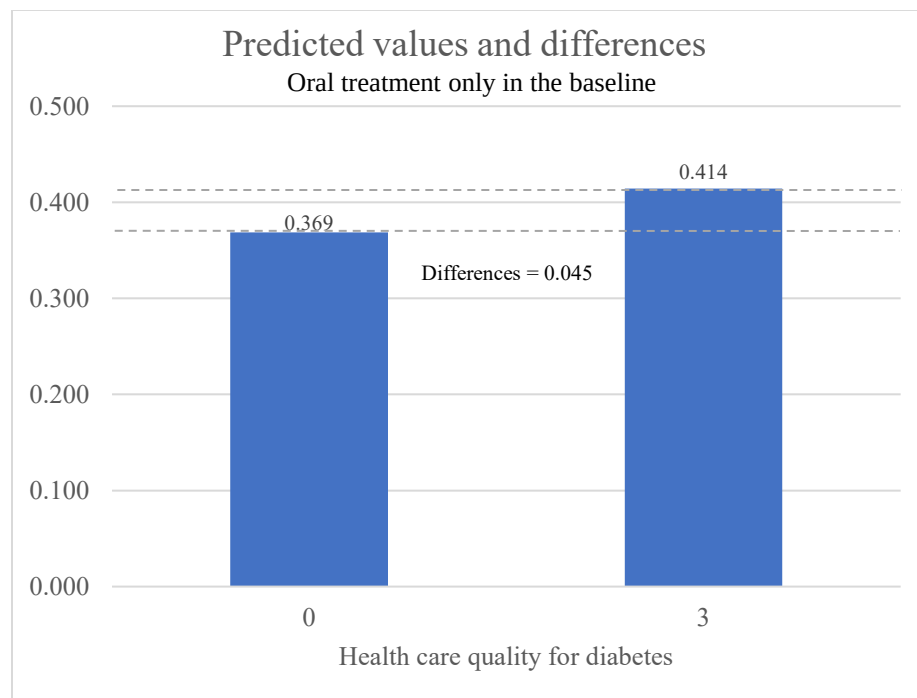
effective management than the average non-Hispanic White American with oral treatment only in the baseline, respectively.

The individual-level variable that affects diabetes care across all treatment regimens is A1c values in the baseline. Higher blood sugar levels in the baseline are associated with a lower likelihood of having effective management in the follow-up. I found that Medicare beneficiaries with severe diabetes in the baseline have 90.7% higher odds of having effective diabetes management than people without Medicare¹³ in the baseline. Medicare beneficiaries receive

¹³ There are only certain times when people can enroll in Medicare. Initial enrollment period begins 3 months before the month s/he turns 65 and ends 3 months after the month s/he turn 65. Most American enroll in Medicare Part A when they turn 65. However, some people consider delaying Medicare Part A until a later date, such as people who

coverage of insulin through Medicare Part B and Part D. The Part B plan covers insulin pumps and related products, whereas the Part D plan covers outpatient insulin products and supplies such as syringes, insulin pens, inhaled insulin devices, and etc. (Hasche et al., 2017). Having Medicare reduces total out-of-pocket health care costs, and Medicare beneficiaries with severe diabetes who need to take insulin are more likely to manage their diabetes effectively than non-Medicare people with severe diabetes.

Figure 1.3. Predicted values and difference in the means to measure the effects of health care quality in regression analysis, stratified by treatment regimens in the baseline



The county-level proxy that captures health care quality is a statistically significant predictor to estimate diabetes management for individuals with oral treatment only in the baseline, the majority group of the sample. The result implies that if the quality of health care for diabetes improves from the worst quality (index value of 0) to the best quality (value of 3), the

contribute to a Health Saving Account or those who have to pay a premium for Part A. In addition, people who have health insurance from their (or their spouse's) current employer may delay enrolling in the Part B plan.

probability of having effective management increases by percentage points (from 0.369 to 0.414), or 10.95% (Figure 1.3). Diabetes is a lifelong disease and needs support from health care practitioners. Affordability and accessibility to adequate treatment is critical for effective diabetes management (Ofstedal et al., 2010; Thorne and Paterson, 2001; Macdonald et al., 2008), leading to low A1c values in the follow-up.

1.5.3. Term effects on effective diabetes management

As noted previously, I identified two groups of continuously enrolled people with diabetes: (1) self-reported individuals with diabetes in the baseline (2006/2008) and (2) people who had no diabetes diagnosis in the baseline but for whom diabetes was newly diagnosed in wave 2 (2010/2012). Thus, the study cohorts consist of 2,521 individuals with known diabetes and 285 individuals with newly diagnosed diabetes. The distinction of two distinct cohorts allowed me to separate the effect of changes in diabetes management over time from possible changes in the people with diabetes over time. Specifically, racial, ethnic, and geographic differences among people with previously diagnosed diabetes may reflect prior differences in the diabetes care. The further distinction of medium or long term among people with previously diagnosed diabetes allowed me to examine whether any individual-level correlates remained constant throughout the study period.

To distinguish the effects of the variables on diabetes management in the medium- and the long-run, I focused on 1,277 individuals¹⁴ with previously diagnosed diabetes who reported their year of diagnosis, which they reported in either the 1992, 2004, or 2006 rounds of data

¹⁴ Among 1,277 unique individuals with previously diagnosed diabetes who reported their year of diagnosis, 887 people with known diabetes appeared twice in the survey, and 780 people with diabetes appeared three times in the survey. Total number of observations is thus 1,677 (887 unique individuals with known diabetes who were responded twice and 390 unique individuals with known diabetes who were responded three times in the survey.). I

collection. Unfortunately, half of individuals with previously diagnosed diabetes in the baseline had missing information for diagnosis year. I categorized the individuals who reported the year of diagnosis into two groups based on diabetes duration (≥ 8 years as long-term and 4-8 years as medium-term). Thus, among 1,277 individuals with previously diagnosed diabetes and who reported the year of diagnosis, 523 unique individuals are in the medium-term group and 754 unique individuals are in the long-term group. To examine the differences between the samples of those who reported their year of diagnosis and those who did not report theirs, I took the baseline data among people with previously diagnosed diabetes and compared the two samples. Non-Hispanic White Americans with known diabetes were more likely to provide their diagnosis year than minority populations with previously diagnosed diabetes. Residents in large central/fringe counties are more likely to provide their year of diagnosis than residents in medium/small metro counties or nonmetro counties (Table A.2). A higher proportion of minority populations with diabetes and rural residents with diabetes may have a long history of diabetes in the baseline because their diabetes was not diagnosed early. This nonrandom selection thus may underestimate racial/ethnic and metro-nonmetro differences in effective diabetes management.

Table 1.5 presents selected baseline characteristics by cohort. The average baseline A1c was well above 6.5% for people with previously diagnosed diabetes in the long-term and close to 6.5% for people with previously diagnosed diabetes in the medium-term or with newly diagnosed diabetes. Individuals with newly diagnosed diabetes were less likely to be obese, more likely to be overweight, and were less likely to have medical treatment in the baseline than people with previously known diabetes. I found that the probability of having effective diabetes management in the follow-up decreases with the duration since diagnosis.

used both rounds of information when a person with diabetes responded more than twice in the survey and took the advantage of a longitudinal data setting.

Table 1.5. Baseline differences in individual-level demographic factors, SES, and health behaviors, and county-level contextual features among people with diabetes by duration of diagnosis

	Previously diagnosed		Newly diagnosed
	≥8 years (N=965)	4-8 years (N=702)	<4 years (N=285)
Number of unique people with known diabetes	754	523	285
Effective management (A1c<6.5%)	28.70%	52.42%	62.46%
Individual characteristics (Baseline)			
Female	52.44%	57.26%	47.72%
Race/Ethnicity			
NHW	51.40%	63.96%	72.98%
NHB	26.53%	19.94%	15.09%
Hispanic	19.27%	12.68%	8.42%
NHO	2.80%	3.34%	3.51%
<i>Metro-nonmetro status</i>			
Large central/fringe metro county	51.61%	51.85%	50.53%
Medium/small metro county	29.12%	29.91%	25.61%
Rural county	19.27%	18.23%	23.86%
Educational attainment			
Less than HS	29.53%	22.22%	20.00%
High School (HS)	33.37%	34.76%	35.09%
Above HS	37.10%	43.02%	44.91%
Married	66.94%	69.23%	68.07%
Has Medicare	61.66%	51.28%	67.72%
Visited a doctor	95.85%	97.15%	96.49%
<i>Smoking history</i>			
Never smoked	42.38%	46.30%	47.37%
Former smoker	43.83%	40.74%	40.70%
Current smoker	13.78%	12.96%	11.93%
Overweight	30.05%	32.34%	34.04%
Obese	57.20%	56.55%	51.58%
Severity of disease (Baseline)			
A1c level	7.28	6.42	6.37
<i>Treatment</i>			
No treatment	9.53%	34.05%	36.49%
Oral only	56.17%	61.25%	60.70%
Insulin only	14.40%	1.57%	1.05%
Both oral and Insulin	19.90%	31.34%	1.75%
County characteristics (Baseline)			
Quality of health care (index)	1.87	1.88	1.87
HPSA	42.01%	35.04%	37.54%

Abbreviations: NHW = non-Hispanic Whites; NHB = non-Hispanic Blacks; NHO = non-Hispanic other races; HPSA = Health Professional Shortage Area

Table 1.6 displays results of the analyses that estimate different duration effects on diabetes management among cohorts with diagnosed diabetes. The models demonstrate that different factors are at play in predicting effective diabetes management across different duration terms. I found statistically significant evidence that there are longer-term effects of the county types on effective diabetes management among individuals with known diabetes. Non-Hispanic Black people with previously diagnosed diabetes in the long run show a decreased likelihood to have effective diabetes management compared to non-Hispanic White Americans with previously diagnosed diabetes in the long run. In terms of health behaviors in the long run (≥ 8 years), people who were classified as overweight or obese in the baseline were 13.3% or 23.3% less likely to manage their diabetes than an individual with diabetes whose weight was below the overweight threshold.

Net of control variables, A1c levels in the baseline were associated with significantly lower odds of having effective diabetes care in the follow-up, with the largest difference among individuals with previously diagnosed diabetes in the medium-term and the smallest difference among individuals with known diabetes who suffer from disease over 8 years. I also found that the odds of having effective management would decrease with degree of severity among people with previously diagnosed diabetes.

Last but not least, net of control variables, I found that the quality of health care contributes more to diabetes management in the short and medium run than in the long run. The results suggest that if the quality of healthcare for diabetes improves from the worst quality (index value of 0) to the best quality (index value of 3), the probability of having effective management of diabetes is increased by 12.76% in the short run and 32.31% in the medium run (Figure 1.4). The estimation results among individuals with known diabetes imply that the

Table 1.6. Results from logistic regression on diabetes management for diabetes by term differences

	Previously diagnosed				Newly diagnosed	
	(1)	(2)	(3)	(4)	(5)	(6)
	Longer term (≥ 8 yrs) [N=965]		Medium term (4 to 8 yrs) [N=702]		Short term (<4 yrs) [N=285]	
Number of unique individuals	754		523		285	
Dep=1 when an individual with diabetes in the baseline does not exceed 6.5% of the A1C level in the follow-up.						
<i>Rurality (ref=Large central/fringe counties)</i>						
Medium/small counties (Baseline)	0.875** (0.161)	0.870** (0.160)	0.696** (0.147)	0.705* (0.149)	1.129 (0.390)	1.129 (0.392)
Rural counties (Baseline)	0.587** (0.137)	0.597** (0.141)	0.788 (0.196)	0.833 (0.214)	1.224 (0.445)	1.238 (0.467)
<i>Race/Ethnicity</i>						
Hispanics	0.965 (0.250)	0.971 (0.245)	1.213 (0.405)	1.257 (0.424)	3.331 (2.135)	3.504 (2.282)
NHB	0.830 (0.172)	0.844 (0.179)	0.889 (0.226)	0.951 (0.247)	2.038 (0.888)	2.113 (0.937)
Female (ref=male)	1.027 (0.176)	1.029 (0.176)	0.845 (0.162)	0.830 (0.159)	1.142 (0.344)	1.142 (0.344)
Individual-level SES and health behaviors (Baseline)						
<i>Education (ref=HS)</i>						
Less than HS	1.273 (0.276)	1.296 (0.283)	1.009 (0.257)	0.996 (0.254)	0.686 (0.282)	0.677 (0.279)
Above HS	1.180 (0.231)	1.184 (0.233)	0.739 (0.154)	0.727 (0.152)	0.976 (0.324)	0.958 (0.320)
Married	0.768 (0.140)	0.772 (0.141)	0.704* (0.145)	0.689* (0.143)	0.908 (0.286)	0.893 (0.284)
Has Medicare	1.544 (0.418)	1.553 (0.421)	0.758 (0.232)	0.751 (0.229)	0.912 (0.481)	0.904 (0.478)
Visited a doctor	2.397* (1.187)	2.433* (1.208)	0.641 (0.395)	0.605 (0.377)	0.443 (0.390)	0.447 (0.394)

Table 1.6. Results from logistic regression on diabetes management for diabetes by term differences (cont)

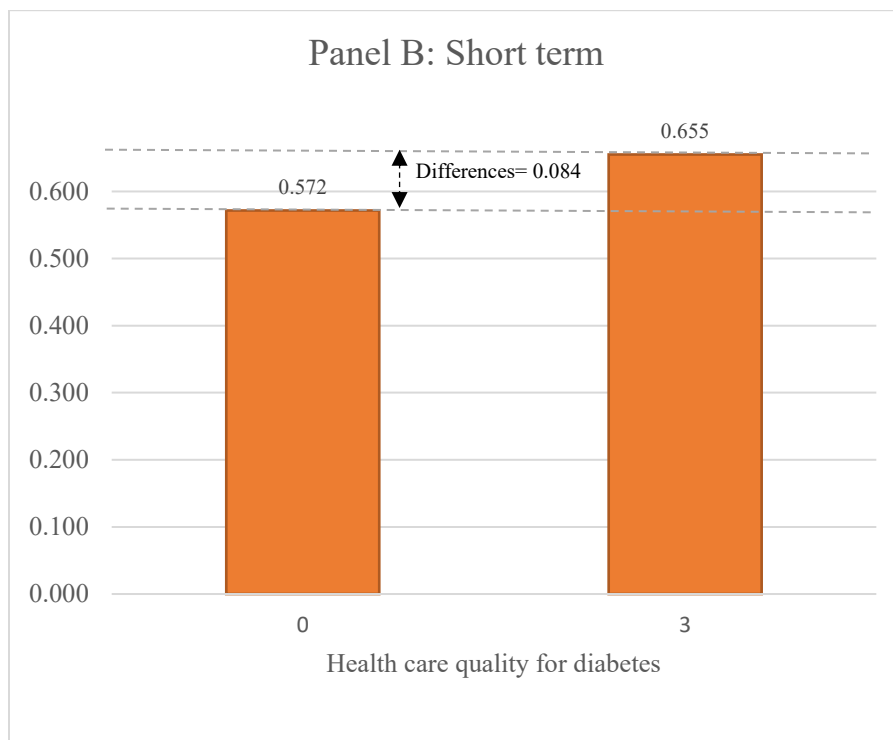
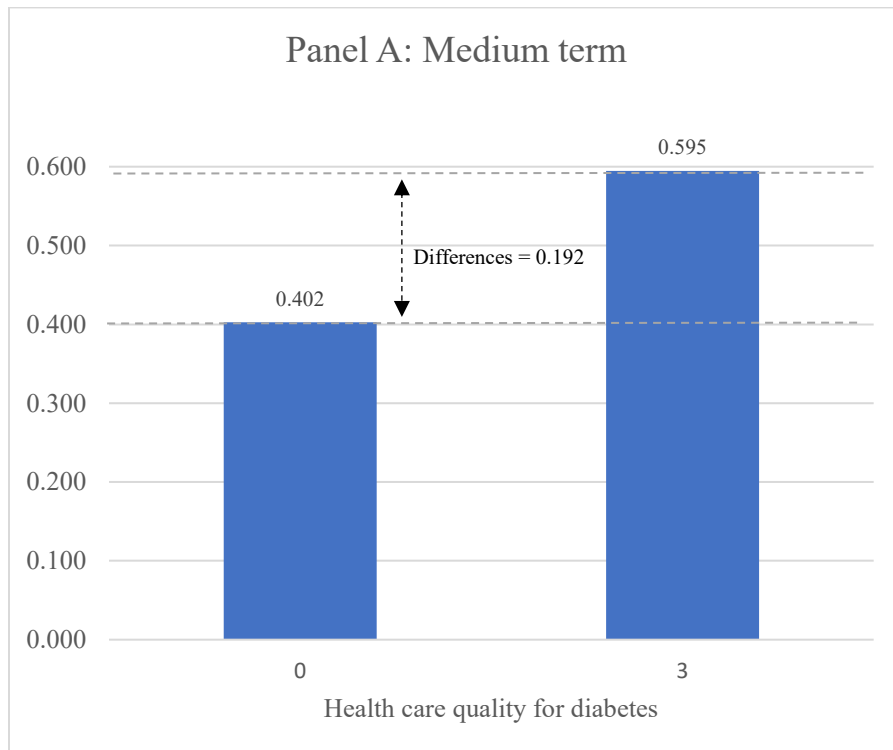
<i>Smoking status (ref=Never smoked)</i>						
Former smoker	1.237 (0.213)	1.235 (0.213)	1.128 (0.218)	1.147 (0.222)	1.132 (0.355)	1.125 (0.353)
Current smoker	1.141 (0.304)	1.141 (0.304)	0.688 (0.204)	0.667 (0.200)	1.813 (0.917)	1.836 (0.940)
Overweight	0.874** (0.225)	0.867** (0.223)	0.881** (0.283)	0.880** (0.284)	0.776 (0.357)	0.784 (0.362)
Being obese	0.766* (0.269)	0.767* (0.269)	0.773** (0.236)	0.775** (0.237)	0.782 (0.354)	0.796 (0.364)
Severity of diabetes (Baseline)						
A1c level	0.599*** (0.046)	0.599*** (0.046)	0.298*** (0.042)	0.297*** (0.041)	0.422*** (0.077)	0.428*** (0.079)
<i>Treatment (ref=None)</i>						
Oral only	0.396*** (0.104)	0.397*** (0.105)	0.572*** (0.110)	0.579*** (0.111)	0.707 (0.214)	0.702 (0.213)
Insulin only	0.231*** (0.080)	0.237*** (0.084)	0.112* (0.130)	0.107* (0.124)	0.145 (0.202)	0.136 (0.190)
Both oral and Insulin	0.263*** (0.084)	0.266*** (0.086)	0.270** (0.166)	0.274** (0.170)	1.943 (2.582)	1.800 (2.396)
County-level contextual factors (Baseline)						
Quality of Health Care		1.173 (0.310)		1.400** (0.348)		1.152* (0.474)
HPSA		0.940 (0.191)		0.924 (0.199)		0.873 (0.301)
Wave FE	X	X	X	X	X	X

Abbreviations: NHW = non-Hispanic Whites; NHB = non-Hispanic Blacks; NHO = non-Hispanic other races; HPSA = Health Professional Shortage Area

Note: All models include dummies for age groups (51-59, 60-64, 65-69, 70-74, and 75 and over), and the four regions in the U.S. Census, with South as the reference category.

Note: Standard errors in parentheses are clustered at the individual level. Significance is indicated by the symbol * at the 10% level, ** at the 5% level, and *** at the 1% level.

Figure 1.4. Predicted values and difference in the means to measure the different term effects of health care quality in regression analysis



quality of health care does not matter if I follow people with diabetes over 8 years, whereas if I follow people with diabetes between 4 and 8 years, the quality of health care matters. As the time since the diagnosis increases, the quality of health care seems to matter less among people with previously diagnosed diabetes. This could be due to measurement error. Quality of health care at a certain time is less indicative of the quality of health care throughout the course of diabetes. Also, what matters is having good quality of care to get diagnosed early and get the right treatment early.

1.5.4. Results with interactions for race/ethnicity, metro-nonmetro status, and gender

It is important to pay attention to both race/ethnicity, metro-nonmetro status, and gender while investigating the effects of each factor on diabetes management. The effects of the degree of rurality may differ across racial/ethnic groups and by gender. Table 1.7 presents the results of the regression including interactions of race/ethnicity with rurality (column 1 and 2) and gender (column 3 and 4), and interactions of rurality and gender (column 5 and 6). The finding that non-Hispanic Black Americans with previously diagnosed diabetes are less likely to manage their diabetes is robust to the inclusion of these interaction terms. In the last two columns, none of the interactions of rurality and gender shows a significant difference from zero.

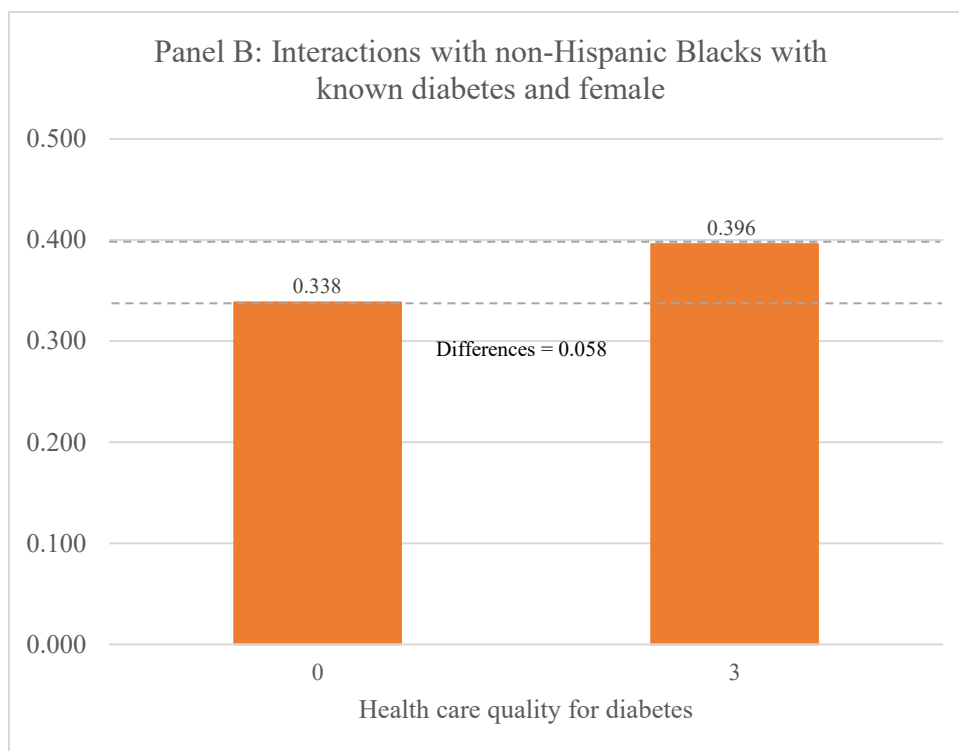
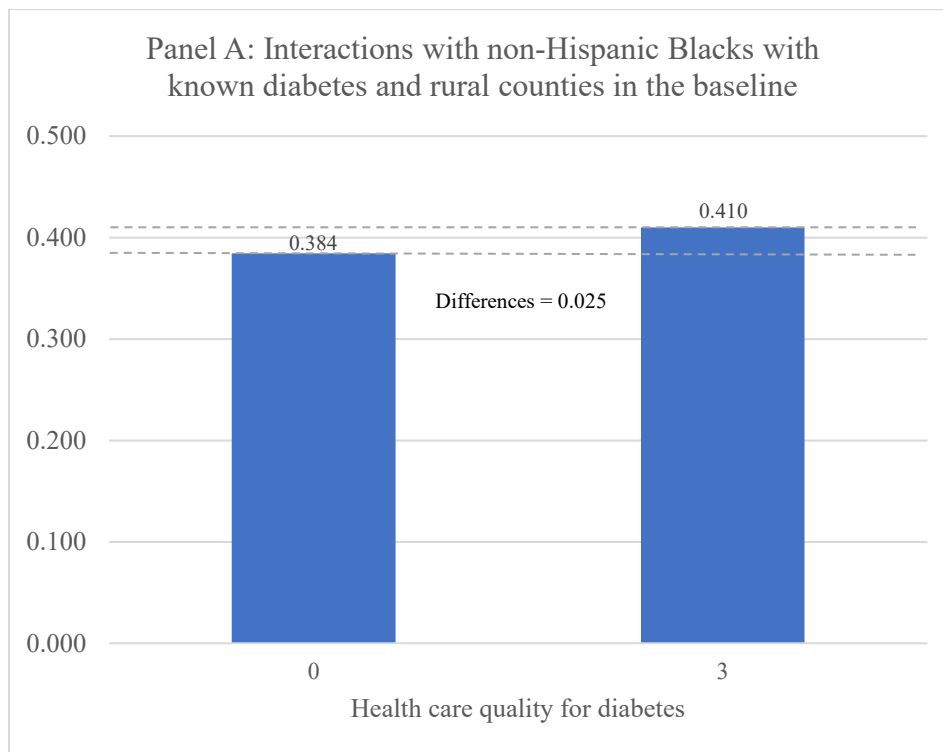
Results for race/ethnicity interacted with metro-nonmetro status (column 1 and 2) show that the average non-Hispanic Black American with previously diagnosed diabetes who lives in a nonmetro county is more likely to manage his/her diabetes than the average non-Hispanic Black American with previously diagnosed diabetes who lives in a metro county. Using non-Hispanic Black Americans with diabetes as the focus variable, the result suggests that being a non-Hispanic Black American with diabetes in a rural county shows 2.54 percentage points (from

Table 1.7. Results with interactions for race/ethnicity, metro-nonmetro status, and gender among individuals with known diabetes (selective coefficients and odds ratios)

	(1) Coefficient	(2) OR	(3) Coefficient	(4) OR	(5) Coefficient	(6) OR
Dep=1 when an individual with diabetes in the baseline does not exceed 6.5% of the A1C level in the follow-up.						
<i>Rurality (ref=Large central/fringe counties)</i>						
Medium/small counties (Baseline)	-0.183	0.833*	-0.216**	0.806**	-0.095	0.91
Rural counties (Baseline)	-0.293**	0.746**	-0.262**	0.770**	-0.246	0.782
<i>Race/Ethnicity</i>						
Hispanics	-0.0696	0.933	-0.109	0.896	-0.217	0.805
NHB	-0.428***	0.652***	-0.403**	0.668**	-0.291***	0.747***
NHO	0.251	1.285	0.233	1.262	-0.018	0.982
Female (ref=male)	0.125	1.133	0.130	1.138	0.201*	1.223*
<i>Interactions with Race/ethnicity and Rurality</i>						
Hispanics x Medium/small counties	-0.357	0.700				
Hispanics x Rural counties	-0.596	0.551				
NHB x Medium/small counties	0.261	1.298				
NHB x Rural counties	0.501**	1.650**				
NHO x Medium/small counties	-0.896*	0.408*				
NHO x Rural counties	-0.311	0.733				
<i>Interactions with Race/ethnicity and Female</i>						
Hispanics x Female			-0.181	0.834		
NHB x Female			0.172**	1.187**		
NHO x Female			-0.446	0.640		
<i>Interactions with Race/ethnicity and Female</i>						
Medium/small counties x Female					-0.225	0.799
Rural counties x Female					-0.036	0.965

Note: All models include dummies for age groups, and the four regions in the U.S. Census, with South as the reference category, and wave effects. Note: Standard errors in parentheses are clustered at the individual level. Significance is indicated by the symbol * at the 10% level, ** at the 5% level, and *** at the 1% level.

Figure 1.5. Predicted values and difference in the means to measure the different effects of interaction terms for race/ethnicity, metro-nonmetro status, and gender among individuals with known diabetes



0.384 to 0.410), or a 6.20 percent higher probability of having effective diabetes management than a non-Hispanic Black American with diabetes in a medium/small metro county (Panel A in Figure 1.5). Results for race/ethnicity interacted with female (column 3 and 4) shows that the average non-Hispanic Black woman with previously diagnosed diabetes is more likely to manage her diabetes than the average non-Hispanic Black man with known diabetes. A change from being a female non-Hispanic Black to being a male non-Hispanic Black increases the probability of having effective diabetes management by 5.77 percentage points (from 0.338 to 0.396), or 14.57 percent (Panel B in Figure 1.5).

1.6. Results for estimation of improvement in severity of diabetes

I have used a cut-off point of 6.5% A1c level to indicate effective management of diabetes. However, if there are significant disparities across race, ethnicity, and metro-nonmetro status in terms of when individuals with diabetes were diagnosed, some people with severe diabetes present ineffective diabetes management. For example, Black Americans with severe diabetes who have A1c levels much over 6.5% in the baseline may have no way to quickly reduce their A1c values to the cut-off in the follow-up. It may overestimate racial and ethnic differences in effective diabetes management. Therefore, I alternatively regressed the change in A1c levels between rounds on the individual-level and contextual county-level factors.

Table 1.8 presents the results for improvement in severity of diabetes measured by the change in A1c levels between rounds. From all four models, I found little evidence that there are effects of the county types, race, and ethnicity on reduction in A1c values between rounds, indicating improvement in severity of diabetes. The effects remain statistically insignificant with the introduction of individual-level SES, health behaviors, disease severity in the baseline, and

Table 1.8. Results of improvement in severity of diabetes among people with previously diagnosed diabetes

	(1)	(2)	(3)	(4)
Dependent variable is the change of A1c level between rounds.				
<i>Rurality (ref=Large central/fringe counties)</i>				
Medium/small counties (Baseline)	-3.510 (5.351)	-2.055 (2.947)	-2.624 (8.105)	-0.346 (5.503)
Rural counties (Baseline)	-9.508 (5.350)	-13.632 (9.990)	-4.624 (12.702)	-8.894 (10.986)
<i>Race/Ethnicity</i>				
Hispanics	0.231 (3.089)	-3.830 (5.206)	-4.247 (14.037)	-0.062 (7.046)
NHB	-2.273 (3.090)	-0.114 (4.325)	-4.966 (5.423)	-0.027 (6.308)
Female (ref=male)	0.168 (2.184)	3.625 (5.430)	3.282 (9.242)	0.304 (5.246)
Individual-level SES and health behavior (Baseline)				
<i>Education (ref=HS)</i>				
Less than HS		3.440 (4.884)	1.934 (2.460)	1.447 (2.664)
Above HS		-2.263 (3.027)	-2.240 (0.436)	-0.250 (3.272)
Married		-1.693 (3.602)	0.789 (2.369)	-0.262 (2.093)
Has Medicare		-1.468 (4.942)	-5.237 (9.467)	-2.056 (4.233)
Visited a doctor		-9.648 (2.627)	5.154 (6.216)	0.175 (6.177)
<i>Smoking status (ref=Never smoked)</i>				
Former smoker		1.194 (2.657)	1.064 (3.671)	-0.008 (3.503)
Current smoker		-1.386 (5.224)	0.742 (8.666)	3.310 (6.272)
Overweight		2.568 (4.845)	-6.398 (3.161)	3.458 (2.451)
Being obese		2.601 (6.178)	-5.108 (2.290)	4.093 (1.891)
Severity of diabetes (Baseline)				
A1c level			-0.404 (0.508)	-0.224 (0.915)

Table 1.8. Results of improvement in severity of diabetes among people with previously diagnosed diabetes (cont)

<i>Treatment (ref=None)</i>				
Oral only			-2.689 (7.617)	-0.319 (3.940)
Insulin only			2.068 (5.731)	3.852 (5.560)
Both oral and Insulin			-0.666 (7.597)	-3.005 (5.200)
County-level contextual factors (Baseline)				
Quality of Health Care				0.201 (2.816)
HPSA				-0.773 (2.554)
Numbers of observation	3,421	3,421	3,421	3,421

Abbreviations: NHW = non-Hispanic Whites; NHB = non-Hispanic Blacks; NHO = non-Hispanic other races; HPSA = Health Professional Shortage Area

Note: All models include dummies for age groups (51-59, 60-64, 65-69, 70-74, and 75 and over), the four regions in the U.S. Census, with South as the reference category, wave fixed effects and individual fixed effects.

Note: Standard errors in parentheses are clustered at the individual level. Significance is indicated by the symbol * at the 10% level, ** at the 5% level, and *** at the 1% level.

county-level contextual variables. Furthermore, I found little evidence that county-level proxies that capture health care infrastructure are important predictors in models for effective diabetes management measured by the difference in A1c levels between rounds. I also estimated improvement in severity of diabetes with the dependent variable of the change in A1c levels between rounds for the regressions by treatment, different term, and analyses with interactions. None of the results for race, ethnicity, and rurality are statistically significant.

1.7. Discussion and conclusion

I investigated the differential effects of race, ethnicity, and rurality in diabetes management using a nationally representative sample of U.S. middle-aged and older adults. This is among the first studies to use a nationally representative sample of U.S. adults in a

longitudinal design that documents racial, ethnic, and metro-nonmetro status differences in diabetes management measured by an objective biomarker. Consistent with recent U.S. research on diabetes care (Kirkbride and Wallace, 2009; Zhu et al., 2011; Sequist et al., 2006; Saydah et al., 2007), this study shows that disparities exist in diabetes management among racial and ethnic groups and across metro-nonmetro status in the United States. I found that individuals with diabetes living in rural counties and non-Hispanic Black Americans with diabetes have the highest odds of having ineffective diabetes management as defined by having A1c levels below the threshold of 6.5%. Previous studies in this area are largely descriptive in nature and stop short of conducting analyses to explain urban-rural differences. In this study, I have built on previous research by using a nationally representative sample of U.S. adults and a multilevel modeling framework that allowed me to examine the extent to which the characteristics of individuals and the health service features of counties attenuate these disparities in diabetes management. A number of important findings emerged.

First, I found that the likelihood of having effective diabetes management decreases with the degree of rurality. The general characteristics for diabetes care in rural America found in this study are consistent with research on a wide range of health measures and structural disadvantages. Rural areas have higher prevalence rates of chronic diseases, self-rated poor health, obesity, and smoking, and experience a lack of adequate access to primary care and specialty care services compared to urban areas (Pathman et al., 2006; Monnat and Pickett, 2011; Glasgow and Brown, 2012). Diabetes management in this study follows a metro-nonmetro gradient. My results demonstrated that as the degree of rurality increases, so does the difference in odds of showing ineffective diabetes management between people living in metropolitan counties and people living in nonmetropolitan counties. The typical resident in a

nonmetropolitan county has greater odds of having poor diabetes management than the typical resident of any of the metropolitan county types, and the disparity becomes greater with the degree of rurality. These differences can be masked when combining all metropolitan counties into one category. Therefore, it is equally important to disaggregate across levels of urban areas as it is to disaggregate across levels of rural areas.

Second, analyses lend support to previous studies investigating racial and ethnic differences in diabetes management among people with diabetes. Specifically, my results present strong and statistically significant racial/ethnic differences for glycemic control with a persistent gap being present between non-Hispanic Black Americans and non-Hispanic White Americans, holding individual- and county-level variables constant. I further found individual-level factors and county-level contextual indicators of health care infrastructure attenuated racial/ethnic disparities in diabetes management. My results are consistent with previous research showing African Americans consistently show worse outcomes and management when compared to other minority populations and non-Hispanic Whites (Campbell et al., 2012; Hausmann et al., 2010), and these findings contribute to related literatures by highlighting the importance of controlling for both individual-level factors and county-level contextual correlates.

Introducing the county-level contextual variables additionally decreases racial/ethnic and metro-nonmetro differences in effective diabetes management. In terms of the county-level contextual factors themselves, health care quality is a salient predictor of effective diabetes management. This finding is consistent with previous research showing public health is highly related to the management of chronic conditions (Epping-Jordan et al., 2004). It further implies that access to specialty health services such as the A1c test, eye examinations, and the blood lipid test is essential for people with diabetes to have effective diabetes management. This finding

held both when I estimated effective diabetes management focusing on people with known diabetes that were diagnosed between 4 and 8 years previously (medium-term) and people with newly diagnosed diabetes (short-term). Thus, it suggests a health care system capable of timely response to a diabetes diagnosis must be considered to achieve effective diabetes management (Mueller and MacKinney, 2006).

I also found non-Hispanic Black Americans with known diabetes who live in metro counties are more likely to manage their diabetes than non-Hispanic Black Americans with known diabetes who live in nonmetro counties. As found in other studies, the average Black female with previously diagnosed diabetes has greater odds of showing ineffective diabetes management than the average non-Hispanic Black male with previously diagnosed diabetes.

Dichotomous threshold measures may not accurately reflect the impact of race, ethnicity, and metro-nonmetro status on diabetes management, and thus I used a continuous measure for A1c level's change between rounds as the other dependent variable. I found little evidence that there are effects of the explanatory variables of interest on the reduction in A1c values between rounds. Aron et al. (2007) showed that continuous weighted A1c measurement based on quality adjusted life years is appropriate in the assessment of quality of diabetes care. However, their sample was from Veterans Health Administration medical centers that may also limit its generalizability. Using an A1c threshold of $\geq 6.5\%$ is recommended to diagnose diabetes and standard in the literature because this value has been shown to be strongly related to complications such as retinopathy (International Expert Committee, 2009). To suggest how health care practitioners and policymakers should deal with the way of optimizing diabetes management, results by looking at the threshold are more informative than outcomes resulted

from the reductions in A1c. Establishing separate diagnostic thresholds based on race, ethnicity, and rurality would be the subject of future research.

Despite the strengths and new contributions of this study, there are some limitations that should be addressed in future research. First, I explore differences in effective diabetes management among middle-aged and old-aged Americans living in three types of counties – large central/fringe metro counties, medium/small metro counties, and rural counties. Combining all of the non-metropolitan counties into the category of rural counties may ignore differences in diabetes care across different types of rural counties. Second, there is heterogeneity within counties that cannot be accounted for in these analyses. For example, many metropolitan counties have outlying rural areas with different resident characteristics compared to those in the urban core of the county. Third, sample sizes are small when I looked at the intersection of gender, race/ethnicity, and rurality. It reduces the power of the analyses and increases the margin of error. Fourth, I cannot tell which types of diabetes an individual with diabetes has and cannot exclude people with Type 1 Diabetes for that reason. As noted above, Type 1 diabetes accounts for 5% of all diagnosed diabetes cases and is diagnosed in childhood¹⁵. Fifth, I could not account for all clinical information for individuals in the sample. Lack of information about duration of medical treatment and duration of the county residence limited my analyses to being stratified by treatment regimens or duration differences. Sixth, minority populations with diabetes were less likely to complete interviews, provide biomarker data, and report their year of diagnosis than non-Hispanic White Americans (Table A.1 and Table A.2). A higher proportion of minority populations with diabetes may have an A1c value over and equal to 6.5% in the baseline and

¹⁵ Among cohorts with diabetes who reported their year of diagnosis for the analyses that estimate different duration effects on effective diabetes management, only 0.71% people were diagnosed under 16 years old. They are in the long-term group (≥ 8 years since diagnosis).

may be diagnosed at late stage, and this nonrandom selection may underestimate racial and ethnic differences in effective diabetes management. Lastly, I could not capture the behavior of rural residents who travel outside their community to obtain advanced health care. Future research should examine whether individuals with diabetes travel outside their counties of residence to have better quality health care services for diabetes.

Taken together, the findings indicate that rural residents with diabetes and non-Hispanic Black Americans with diabetes have the greatest likelihood of having ineffective diabetes management and show that adequate health care service is a vital predictor of diabetes care. Without a clear understanding of the geographic variation and without controlling for county-level contextual features with individual-level characteristics, policy solutions may be misdirected. McLaughlin et al. (2007) pointed out that understanding the combination of individual and contextual characteristics facilitates regional responses and better resource targeting.

In addition, it is important to understand whether these differences come from individual factors or environmental factors. If disparities are led by individual-level correlates, an intervention that improves people's health behaviors is needed. On the other hand, if disparities are caused by environmental issues, distinct policy implications should be given. For example, future studies may investigate county-level food access that might explain differences in effective diabetes management across race, ethnicity, and metro-nonmetro status. Therefore, future public health efforts designed to reduce racial, ethnic, and metro-nonmetro differences in diabetes management should examine detailed urban-rural classifications while controlling for individual characteristics, health behaviors, diabetes severity, and county-level contextual correlates.

1.8. Chapter 1 references

- Aron, David, Rajan, Mangala, and Leonard M. Pogach. 2007. "Summary measures of quality of diabetes care: comparison of continuous weighted performance measurement and dichotomous thresholds." *International Journal for Quality in Health Care*, 19(1): 29-36.
- Aspray, Terence J., and Nigel Unwin. 2006. "Clinical guidelines for older adults with diabetes mellitus." *JAMA* 296(15): 1839-1840.
- American Diabetes Association. 2018. "Economic costs of diabetes in the US in 2017." *Diabetes care* 41(5): 917-928.
- Ailshire, Jennifer, Sarah Mawhorter, and Eun Young Choi. 2020. "Contextual Data Resource (CDR): US Decennial Census and American Community Survey Data, 1990-2018" Version 2.0. Los Angeles, CA: USC/UCLA Center on Biodemography and Population Health.
- Basu, Jayasree, Bernard Friedman, and Helen Burstin. 2004. "Managed care and preventable hospitalization among Medicaid adults." *Health Services Research* 39(3):489-510.
- Bray, George A. 1992. "Drug treatment of obesity." *The American journal of clinical nutrition* 55(2): 538S-544S.
- Brown, Arleen F., Edward W. Gregg, Mark R. Stevens, Andrew J. Karter, Morris Weinberger, Monika M. Safford, Tiffany L. Gary et al. 2005. "Race, ethnicity, socioeconomic position, and quality of care for adults with diabetes enrolled in managed care: the Translating Research Into Action for Diabetes (TRIAD) study." *Diabetes care* 28(12): 2864-2870.
- Brown, A.F., Mangione, C.M., Saliba, D. and Sarkisian, C.A., 2003. "California Healthcare Foundation/American Geriatrics Society Panel on Improving Care for Elders with Diabetes Guidelines for improving the care of the older person with diabetes mellitus." *J Am Geriatr Soc* 51(5): S265-80.
- Buell, Catherine, Duclie Kermah, and Mayer Davidson. 2007. "Utility of A1C for diabetes screening in the 1999–2004 NHANES population." *Diabetes care* 30(9): 2233-2235.
- Casey, Michelle, Kathleen Thiede Call, and Jill Klingner. 2000. "The influence of rural residence on the use of preventive health care services." Division of Health Services Research and Policy, School of Public Health, University of Minnesota.
- Campbell, Jennifer A., Rebekah J. Walker, Brittany L. Smalls, and Leonard E. Egede. 2012. "Glucose control in diabetes: the impact of racial differences on monitoring and outcomes." *Endocrine* 42(3): 471-482.
- Chatterji, Pinka, Heesoo Joo, and Kajal Lahiri. 2012. "Beware of being unaware: racial/ethnic

- disparities in chronic illness in the USA.” *Health economics* 21(9): 1040-1060.
- Cheng, Yiling J., Alka M. Kanaya, Maria Rosario G. Araneta, Sharon H. Saydah, Henry S. Kahn, Edward W. Gregg, Wilfred Y. Fujimoto, and Giuseppina Imperatore. 2019. “Prevalence of diabetes by race and ethnicity in the United States, 2011-2016.” *Jama* 322(24): 2389-2398.
- Chew, Lisa D., Dean Schillinger, Charles Maynard, Daniel S. Lessler, and Consortium for Quality Improvement in Safety Net Hospitals. 2008. “Glycemic and lipid control among patients with diabetes at six US public hospitals.” *Journal of Health Care for the Poor and Underserved* 19(4): 1060-1075.
- Chung, Hyeran and Mary Arends-Kuenning. 2022. “Racial, ethnic, and urban/rural differences in the transition into diabetes: Evidence from the Health and Retirement Survey biomarker and self-reported data” *Journal of Rural Social Sciences* 37(1): Article 5.
<https://egrove.olemiss.edu/jrss/vol37/iss1/5>
- Crimmins, Eileen, Jessica Faul, Jung Ki Kim, Heidi Guyer, Kenneth Langa, Mary Beth Ofstedal, Amanda Sonnega, Robert Wallace, and David Weir. 2013. *Documentation of Biomarkers in the 2006 and 2008 Health and Retirement Study*. Ann Arbor, MI: Institute for Social Research.
- Danaei, Goodarz, Ari B. Friedman, Shefali Oza, Christopher JL Murray, and Majid Ezzati. 2009. “Diabetes prevalence and diagnosis in US states: analysis of health surveys.” *Population health metrics* 7(1): 1-13.
- Durso, Samuel C. 2006. “Using clinical guidelines designed for older adults with diabetes mellitus and complex health status.” *JaMa* 295(16): 1935-1940.
- Epping-Jordan, J. E., S. D. Pruitt, R. Bengoa, and Edward H. Wagner. 2004. “Improving the quality of health care for chronic conditions.” *BMJ Quality & Safety* 13(4): 299-305.
- Gallagher, Emily, Derek Le Roith, and Zachary Bloomgarden. 2009. “Review of hemoglobin A1c in the management of diabetes.” *Journal of diabetes* 1(1): 9-17.
- Ginde, Adit, Enrico Cagliero, David Nathan, and Carlos Camargo. 2008. “Value of risk stratification to increase the predictive validity of HbA1c in screening for undiagnosed diabetes in the US population.” *Journal of general internal medicine* 23(9): 1346-1353.
- Glasgow, Nina, and David L. Brown. 2012. “Rural ageing in the United States: Trends and contexts.” *Journal of Rural Studies* 28(4): 422-431.
- Gong, Gordon, Scott G. Phillips, Catherine Hudson, Debra Curti, and Billy U. Philips. 2019. “Higher US rural mortality rates linked to socioeconomic status, physician shortages, and lack of health insurance.” *Health Affairs* 38(12): 2003-2010.

- Hartley, David. 2004. "Rural health disparities, population health, and rural culture." *American journal of public health* 94(10): 1675-1678.
- Hasche, Jennifer, Christopher Ward, and Nicholas Schluterman. 2017. "Diabetes occurrence, costs, and access to care among medicare beneficiaries aged 65 years and over." *Centers for Medicare and Medicaid Services Office of Enterprise Data & Analytics*.
- Hausmann, Leslie RM, Dianxu Ren, and Mary Ann Sevick. 2010. "Racial differences in diabetes-related psychosocial factors and glycemic control in patients with type 2 diabetes." *Patient preference and adherence* 4: 291-299.
- International Expert Committee. 2009. "International Expert Committee report on the role of the A1C assay in the diagnosis of diabetes." *Diabetes care* 32(7): 1327-1334.
- Inter-university Consortium for Political and Social Research (ICPSR). 2008. *County Characteristics, 2000-2007* [United States]. Ann Arbor, MI: Inter-university Consortium for Political and Social Research. <https://doi.org/10.3886/ICPSR20660.v2>
- James, Wesley L. 2014. "All rural places are not created equal: revisiting the rural mortality penalty in the United States." *American journal of public health* 104(11): 2122-2129.
- Johnston, Kenton J., Hefei Wen, and Karen E. Joynt Maddox. 2019. "Lack of access to specialists associated with mortality and preventable hospitalizations of rural Medicare beneficiaries." *Health Affairs* 38(12): 1993-2002.
- Juster, F. Thomas and Richard Suzman. 1995. "An Overview of the Health and Retirement Study." *Journal of Human Resources* 30(supplement): S7-S56. <https://doi.org/10.2307/146277>
- Kirkbride, Kelly, and Neal Wallace. 2009. "Rural Health Clinics and Diabetes-Related Primary Care for Medicaid Beneficiaries in Oregon." *The Journal of Rural Health* 25(3): 247-252.
- Kirkman, M. Sue, Vanessa Jones Briscoe, Nathaniel Clark, Hermes Florez, Linda B. Haas, Jeffrey B. Halter, Elbert S. Huang et al. 2012. "Diabetes in older adults." *Diabetes care* 35(12): 2650-2664.
- Krishna, Santosh, Kathleen N. Gillespie, and Timothy M. McBride. 2010. "Diabetes burden and access to preventive care in the rural United States." *The Journal of Rural Health* 26(1): 3-11.
- Levy, Leonard. 2021. "The Exploding Incidence and Prevalence of Diabetes." *The Diabetic Foot* 151-155.
- Lutfiyya, M. Nawal, Yogi R. Patel, John B. Steele, Beatrice S. Tetteh, Linda Chang, Carlos

- Aguero, Om Prakash, and Martin S. Lipsky. 2009. "Are there disparities in diabetes care? A comparison of care received by US rural and non-rural adults with diabetes." *Primary Health Care Research & Development* 10(4): 320-331.
- Lutfiyya, M. Nawal, Joel E. McCullough, Lori Mitchell, L. Scott Dean, and Martin S. Lipsky. 2011. "Adequacy of diabetes care for older US rural adults: a cross-sectional population-based study using 2009 BRFSS data." *BMC Public Health* 11(1): 1-11.
- Macdonald, Wendy, Anne Rogers, Tom Blakeman, and Peter Bower. 2008. "Practice nurses and the facilitation of self-management in primary care." *Journal of advanced nursing* 62(2): 191-199.
- Machado, Sara R., Sahan Jayawardana, Elias Mossialos, and Muthiah Vaduganathan. 2021. "Physician density by specialty type in urban and rural counties in the US, 2010 to 2017." *JAMA network open* 4(1): e2033994-e2033994.
- Mainous, Arch G., Dana E. King, David R. Garr, and William S. Pearson. 2004. "Race, rural residence, and control of diabetes and hypertension." *The Annals of Family Medicine* 2(6): 563-568.
- McConnel, Charles E., and Marion R. Zetzman. 1993. "Urban/rural differences in health service utilization by elderly persons in the United States." *The Journal of Rural Health* 9(4): 270-280.
- McLaughlin, Diane, Shannon Stokes, and Atsuko Nonoyama. 2001. "Residence and Income Inequality: Effects on Mortality Among U.S. Counties." *Rural Sociology* 66(4):579-598.
- McLaughlin, Diane, Shannon Stokes, Johnelle Smith, and Atsuko Nonoyama. 2007. "Differential Mortality Across the United States. The influence of Place-Based Inequality." *The Sociology of Spatial Inequality, State University of New York Press, Albany, NY*: 141-162.
- McWilliams, J. Michael, Ellen Meara, Alan M. Zaslavsky, and John Z. Ayanian. 2009. "Differences in control of cardiovascular disease and diabetes by race, ethnicity, and education: US trends from 1999 to 2006 and effects of Medicare coverage." *Annals of internal medicine* 150(8): 505-515.
- Mercado, Carla I., Kai McKeever Bullard, Edward W. Gregg, Mohammed K. Ali, Sharon H. Saydah, and Giuseppina Imperatore. 2021. "Differences in US Rural-Urban Trends in Diabetes ABCS, 1999–2018." *Diabetes Care*. <https://doi.org/10.2337/dc20-0097>
- Meneilly, Graydon S., and Daniel Tessier. 2001. "Diabetes in elderly adults." *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences* 56(1): M5-M13.

- Monnat, Shannon, and Camille Beeler Pickett. 2011. "Rural/urban differences in self-rated health: Examining the roles of county size and metropolitan adjacency." *Health & Place* 17(1): 311-319. <https://doi.org/10.1016/j.healthplace.2010.11.008>
- Mueller, Keith J., and A. Clinton MacKinney. 2006. "Care across the continuum: access to health care services in rural America." *Journal of rural health* 22(1): 43-49.
- Nakanishi, Noriyuki, Koji Nakamura, Yoshio Matsuo, Kenji Suzuki, and Kozo Tatara. 2000. "Cigarette smoking and risk for impaired fasting glucose and type 2 diabetes in middle-aged Japanese men." *Annals of internal medicine* 133(3): 183-191.
- Narayan, K. Venkat, James P. Boyle, Linda S. Geiss, Jinan B. Saaddine, and Theodore J. Thompson. 2006. "Impact of recent increase in incidence on future diabetes burden: US, 2005–2050." *Diabetes care*, 29(9): 2114-2116.
- Oba, Shino, Mitsuhiro Noda, Kayo Waki, Akiko Nanri, Masayuki Kato, Yoshihiko Takahashi, Kalpana Poudel-Tandukar et al. 2012. "Smoking cessation increases short-term risk of type 2 diabetes irrespective of weight gain: the Japan Public Health Center-Based Prospective Study." *PloS one* 7(2): e17061.
- Oftedal, Bjørg, Bjørg Karlsen, and Edvin Bru. 2010. "Perceived support from healthcare practitioners among adults with type 2 diabetes." *Journal of advanced nursing* 66(7): 1500-1509.
- Ong, Kwok Leung, Bernard MY Cheung, Yu Bun Man, Chu Pak Lau, and Karen SL Lam. 2007. "Prevalence, awareness, treatment, and control of hypertension among United States adults 1999–2004." *Hypertension* 49(1): 69-75.
- O'Connor, Alane, and Gregory Wellenius. 2012. "Rural–urban disparities in the prevalence of diabetes and coronary heart disease." *Public health* 126(10): 813-820.
- Panelli, Ruth, Lou Gallagher, and Robin Kearns. 2006. "Access to rural health services: Research as community action and policy critique." *Social Science & Medicine* 62(5): 1103-1114.
- Pathman, Donald E., Thomas C. Ricketts III, and Thomas R. Konrad. 2006. "How adults' access to outpatient physician services relates to the local supply of primary care physicians in the rural southeast." *Health services research* 41(1): 79-102.
- Resnick, Helaine E., Janice Heineman, Robyn Stone, and Ronald I. Shorr. 2008. "Diabetes in US nursing homes, 2004." *Diabetes care* 31(2): 287-288.
- Rosenblatt, Roger A., C. Holly A. Andrilla, Thomas Curtin, and L. Gary Hart. 2006. "Shortages of medical personnel at community health centers: implications for planned expansion." *Jama* 295(9): 1042-1049.

- Saaddine, Jinan, Michael Engelgau, Gloria Beckles, Edward Gregg, Theodore Thompson, and KM Venkat Narayan. 2002. "A diabetes report card for the United States: quality of care in the 1990s." *Annals of internal medicine* 136(8): 565-574.
- Saydah, Sharon, Catherine Cowie, Mark S. Eberhardt, Nathalie De Rekeneire, and KM Venkat Narayan. 2007. "Race and ethnic differences in glycemic control among adults with diagnosed diabetes in the United States." *Ethnicity & disease* 17(3): 529-535.
- Scannell, Christopher A., Jacob K. Quinton, Nicholas J. Jackson, and Yusuke Tsugawa. 2021. "Primary Care Health Professional Shortage Area Designations Before and After the Affordable Care Act's Shortage Designation Modernization Project." *JAMA network open* 4(7): e2118836-e2118836.
- Sequist, Thomas D., Alyce Adams, Fang Zhang, Dennis Ross-Degnan, and John Z. Ayanian. 2006. "Effect of quality improvement on racial disparities in diabetes care." *Archives of Internal Medicine* 166(6): 675-681.
- Singh, Gopal K., Gem P. Daus, Michelle Allender, Christine T. Ramey, Elijah K. Martin, Chris Perry, Andrew A. De Los Reyes, and Ivy P. Vedamuthu. 2017. "Social determinants of health in the United States: addressing major health inequality trends for the nation, 1935-2016." *International Journal of MCH and AIDS* 6(2): 139-164.
- Singh, Gopal K., Mohammad Siahpush, Romuladus E. Azuine, and Shanita D. Williams. 2015. "Widening socioeconomic and racial disparities in cardiovascular disease mortality in the United States, 1969-2013." *International Journal of MCH and AIDS* 3(2): 106-118.
- Silverman, Robert A., Harman Arora, Thomas M. Kwiatkowski, Kirsten K. Graff, Emily Baneman, Nina Kohn, Cynthia Foss Bowman, Kelly N. Smith, and Martin L. Lesser. 2011. "Increased A1c among adult emergency department patients with type 2 diabetes." *Annals of emergency medicine* 57(6): 575-581.
- Smalls, Brittany L., Tiarney D. Ritchwood, Kinf G. Bishu, and Leonard E. Egede. 2020. "Racial/ethnic differences in glycemic control in older adults with type 2 diabetes: United States 2003–2014." *International journal of environmental research and public health* 17(3): 950.
- Thackeray, Rosemary, Ray M. Merrill, and Brad L. Neiger. 2004. "Disparities in diabetes management practice between racial and ethnic groups in the United States." *The Diabetes Educator* 30(4): 665-675.
- Thorne, Sally E., and Barbara L. Paterson. 2001. "Health care professional support for self-care management in chronic illness: insights from diabetes research." *Patient education and counseling* 42(1): 81-90.

- UK Prospective Diabetes Study (UKPDS) Group. 1998. "Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33)." *The lancet* 352(9131): 837-853.
- Voulgari, Christina, Nicholas Katsilambros, and Nicholas Tentolouris. 2011. "Smoking cessation predicts amelioration of microalbuminuria in newly diagnosed type 2 diabetes mellitus: a 1-year prospective study." *Metabolism* 60(10): 1456-1464.
- Weir, David. 2008. "Elastic powers: The integration of biomarkers into the Health and Retirement Study." *Biosocial Surveys*. National Academies Press (US).
- Yeh, Hsin-Chieh, Bruce B. Duncan, Maria Inês Schmidt, Nae-Yuh Wang, and Frederick L. Brancati. 2010. "Smoking, smoking cessation, and risk for type 2 diabetes mellitus: a cohort study." *Annals of internal medicine* 152(1): 10-17.
- Zhang, Xuanping, Linda S. Geiss, Yiling J. Cheng, Gloria L. Beckles, Edward W. Gregg, and Henry S. Kahn. 2008. "The missed patient with diabetes: how access to health care affects the detection of diabetes." *Diabetes care* 31(9): 1748-1753.
- Zhao, Guixiang, Earl S. Ford, Chaoyang Li, and Lina S. Balluz. 2011. "Physical activity in US older adults with diabetes mellitus: prevalence and correlates of meeting physical activity recommendations." *Journal of the American Geriatrics Society* 59(1): 132-137.
- Zhu, Vivienne J., Wanzhu Tu, David G. Marrero, Marc B. Rosenman, and J. Marc Overhage. 2011. "Race and medication adherence and glycemic control: findings from an operational health information exchange." *AMIA Annual Symposium Proceedings*, American Medical Informatics Association 2011: 1649-1657.

CHAPTER 2

RACIAL, ETHNIC, AND URBAN/RURAL DIFFERENCES IN TRANSITIONS INTO

DIABETES: EVIDENCE FROM THE HEALTH AND RETIREMENT SURVEY

BIOMARKER AND SELF-REPORTED DATA

with Mary Arends-Kuenning

2.1 Introduction

The prevalence of diabetes has been increasing in the United States. Between 1990 and 2015, the number of people living with diagnosed diabetes more than tripled from 6.21 million to 23.35 million (Centers for Disease Control and Prevention, 2018), and the incidence rates of diagnosed diabetes range from 3.3 cases per 1,000 in 1980 to 7.8 cases per 1,000 in 2007 (Boyle et al. 2010). Diabetes is a disease that occurs when blood sugar is too high in the bloodstream. In people with diabetes, the body either does not produce enough insulin, which is a hormone that controls blood sugar, or the body cannot use the insulin it produces effectively. When insulin levels are too low or cells stop responding to insulin, too much blood sugar stays in bloodstream, which can cause serious health problems including heart disease, vision loss, kidney disease, and morbidity (Centers for Disease Control and Prevention, 2020). There are two main types of diabetes¹⁶. Type 1 diabetes occurs most frequently in children and accounts for less than 10 percent of all diabetes cases (Mobasserri et al., 2020). Type 2 diabetes is more common in adults and is the focus of this study. In type 2 diabetes, cells are not able to utilize the insulin produced by the body. Treatment includes lifestyle changes, metformin taken orally, insulin injections, or

¹⁶ Gestational diabetes is one of types of diabetes that consists of high blood glucose during pregnancy. However, it usually disappears after pregnancy. Mothers and their children may be at increased risk of developing type 2 diabetes later in life (Zhu and Zhang, 2016).

combination therapy, and the recommended treatment depends on the stage. Stage 1 represents “insulin resistance,” Stage 2 “prediabetes,” Stage 3 “type 2 diabetes,” and Stage 4 “type 2 diabetes with complications.” For those diagnosed with Stage 1 diabetes, medical advice focuses on lifestyle adjustments through diet and exercise to prevent the disease from progressing (Mechanick et al., 2018).

The burden of diabetes in the U.S. is substantial. For individuals, the psychological and quality-of-life costs are significant (Pearce et al., 2013; Falco et al., 2015). Medical care costs are high and increasing (Brandle et al., 2003; Walker et al., 2020). The economic burden associated with diabetes and prediabetes exceeded \$327 billion in 2017, consisting of \$237 billion in direct medical costs and \$90 billion in reduced productivity. This national estimate is 51 percent higher than the \$218 billion estimate for 2007 (American Diabetes Association, 2018). In addition, diabetes ranked fourth among the causes of age-standardized Disability-Adjusted Life Years (DALY) in the U.S. in 2016, and the DALYs attributed to diabetes in the U.S. increased by 11 percent between 1990 and 2016 (The U.S. Burden of Disease Collaborators, 2018).

Unhealthy lifestyle behaviors such as poor diet and lack of exercise are the major causes driving the increase in diabetes to epidemic proportions. Individual-level characteristics also affect the prevalence of diabetes. As people age, they are more likely to develop diabetes, and the estimated prevalence of diabetes among seniors aged 65 years and over in the U.S. reached 26.8 percent in 2018 (Centers for Disease Control and Prevention, 2020). The other demographic factors are changes in racial and ethnic composition¹⁷. The U.S. Census Bureau’s (2015) projections suggest that over 50 percent of Americans will be nonwhite by 2044. It is well-

¹⁷ Minority populations experience different levels of diabetes prevalence. According to the Centers for Disease Control and Prevention (2020), non-Hispanic White Americans report the lowest rates (11.9 percent), followed by Hispanics (14.7 percent), non-Hispanic Asians (14.9 percent), and non-Hispanic Black Americans (16.4 percent).

documented that racial and ethnic minorities have a higher prevalence of diabetes than non-minority individuals (Golden et al., 2012). Multiple factors contribute to these disparities. Biological factors include insulin resistance, hyperglycemia, visceral adiposity, and obesity, which vary across racial and ethnic groups (Golden et al., 2012). Racial and ethnic minorities experience more stress than non-Hispanic Whites, which leads to higher allostatic loads that are related to diabetes (Golden et al., 2012). Health behaviors such as exercising and self-monitoring of blood glucose also vary by race and ethnicity (Spanakis and Golden, 2013). Racial and ethnic minorities are more likely to live in low-income neighborhoods than non-Hispanic whites, which affects health through ability to exercise, access to healthy foods, and exposure to crime and other stressors. Access to health care also varies across racial and ethnic groups due to access to health insurance and provider locations, among other factors (Spanakis and Golden, 2013).

Residential differences are associated with individual health, above and beyond individual-level characteristics (Duncan et al., 1996; Robert, 1998; Subramanian et al., 2001; Macintyre et al., 2002; Boardman, 2004; Dubowitz et al., 2008; Monnat and Pickett, 2011). National standards consistently recommend that people with diabetes receive care from a multidisciplinary team of physicians, nurses, nutritionists, and exercise experts certified as diabetes educators with knowledge of behavioral psychology (Ceballos et al., 2010). However, residents of rural America experience a relative shortage of physicians (Rosenblatt and Hart, 2000; Eberhardt et al., 2001) and are more likely to have limited access to health care services (Bolen et al., 2000; Waidmann and Rajan, 2000; Mainous et al., 2004) than residents of urban America. Such health delivery challenges of rural America make rural residents less likely to have regular check-ups, thus limiting access to adequate diabetes care and increasing the likelihood of rural people having to seek urgent or emergent diabetes care (Basu et al., 2004). Exercise plays a major role

in the prevention and control of diabetes. Adherence to an exercise program is critical for optimal health in individuals with diabetes (Colberg et al., 2010).

Resources and senior-focused amenities such as a senior club are centered within urban areas (Vogelsang, 2016), and rural people are less likely to have opportunities to participate in group exercise than urban people. However, the extent to which these structural differences exist for transitions between stages of type 2 diabetes is unclear because prior research in this area has been restricted to a certain state or geographic region (Goonesekera et al., 2015) or focused on the population with diabetes at a fixed point in time.

As James and her coauthors (2017) show, people of color in rural communities generally suffer worse health than non-Hispanic Whites in those communities. When examining racial and ethnic disparities in a rural context, it is important to recognize the specific kinds of resilience that rural communities may confer. Different locales expose individuals to different regional historical, policy, environmental, and social contexts. For example, rural non-Hispanic Black adults are much more likely to live in the South, with legacies of segregation.

The aim of this analysis is to assess whether there are associations among the characteristics of rurality and race/ethnicity and the outcome of the development of diabetes in older adults and to what extent individual- and county-level determinants explain transitions along the continuum of rurality. The questions are addressed with data from the Health and Retirement Study that is a national survey of U.S. adults above the age of 50 years, the 1990-2000 U.S. Censuses, the Dartmouth Atlas of Health Care, and the Inter-university Consortium for Political and Social Research. We examine separately the transition from prediabetes to diabetes, and the transitions from having no diabetes to prediabetes and diabetes.

2.2 Background

2.2.1. *Contribution of this study*

This paper extends previous research on racial, ethnic, and urban-rural differences in diabetes in four ways. First, we use both biomarker data and self-reported doctors' diagnoses of diabetes, which allow us to identify people with prediabetes as well as people with undiagnosed diabetes. Estimates based only on self-reports of physicians' diagnoses may underestimate diabetes (Golden et al., 2012). Second, we focus on type 2 diabetes by employing logistic models with national data representing U.S. adults over 50 years of age. Third, we assess whether individual characteristics and county-level contextual features explain the urban-rural differences in the development of diabetes. Fourth, we use longitudinal data to focus on individual transitions from prediabetes into diabetes and from a normal state into prediabetes and diabetes. This paper complements previous research and suggests directions for future research.

2.2.2. *Hypotheses of this study*

We hypothesize that the degree of rurality will be associated with transitions between stages of type 2 diabetes in the following way. The average individual living in a nonmetropolitan county will have greater odds of reporting the development of diabetes than the average individual living in a metropolitan county, and this difference in odds will increase as rurality increases. We also hypothesize the associations with rurality will be attenuated when controlling for additional contextual factors.

2.3 Data and Descriptive statistics

2.3.1. Data

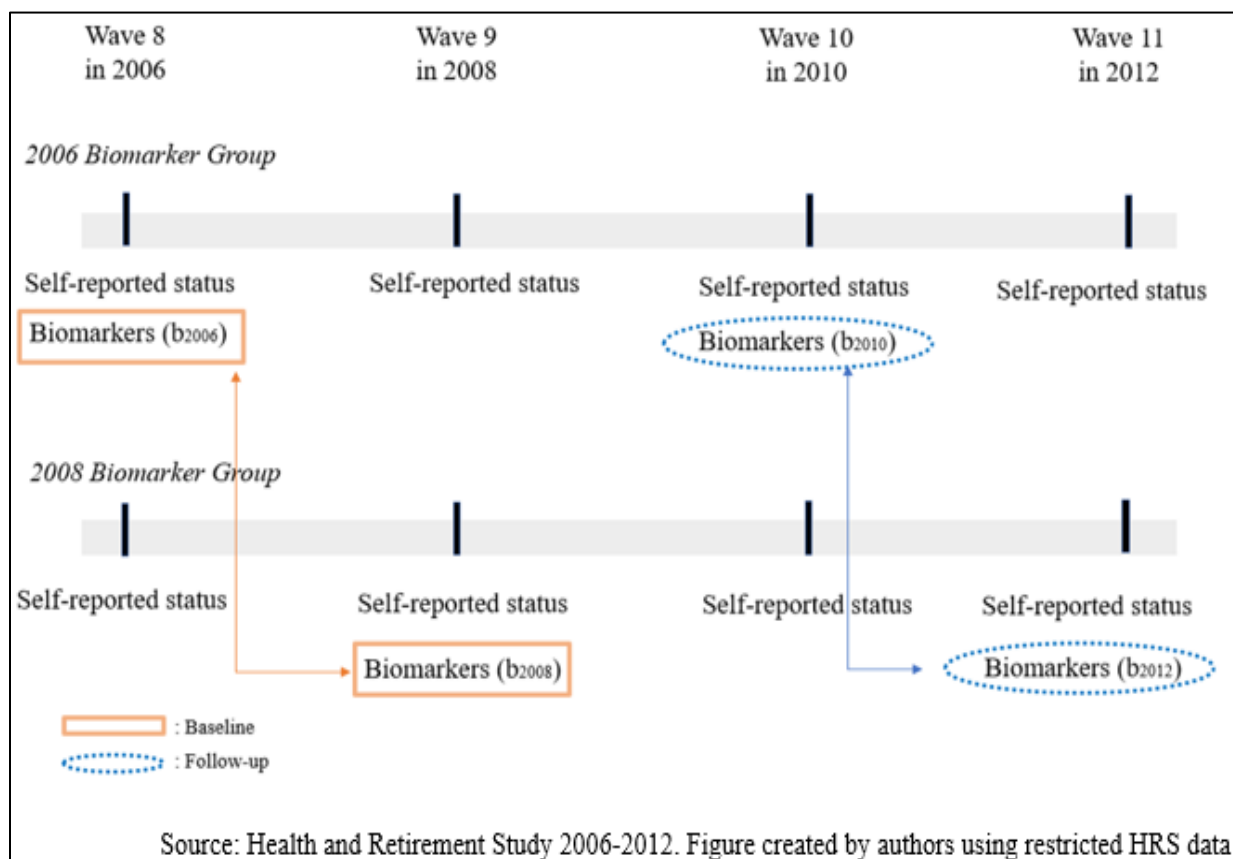
Individual-level data are taken from the 2006-2012 Health and Retirement Study. The Health and Retirement Study (HRS) is a biennial panel survey of U.S. households sponsored by the National Institute on Aging and conducted by the University of Michigan's Institute for Social Research (Juster and Suzman, 1995). Originally begun in 1992 with a representative sample of Americans born between 1931 and 1941, the HRS is representative of the entire U.S. population over 50 years of age with a national sample of over 30,000 individuals.

Respondents are surveyed for self-reported outcomes biennially, whereas the HRS expanded to include biomarkers in Enhanced Face-to-Face (EFTF) interviews conducted on randomly selected rotating halves of the panel in each wave starting in 2006. The sample was selected at the household level, and each selected household was required to include at least one age-eligible member. To examine the differences between the samples of those who decided to provide biomarkers and those who decided against it, we took the baseline data and compared the samples. Non-Hispanic White Americans were more likely to complete the interview and provide the biological samples than non-Hispanic Black Americans or Hispanics, and people living in metro counties were more likely to provide biological samples than people living in nonmetro counties.

During the interviews, researchers measured and collected two types of biological samples to evaluate biomarkers: blood and saliva. Saliva is for DNA extraction, and blood is used to measure 5 biomarkers, which are total cholesterol, High-Density-Lipoprotein (HDL) cholesterol, Glycated hemoglobin (HbA1c, an indicator of glycemic control over the past 2-3 months), C-reactive protein, and Cystatin C (Crimmins et al., 2013). Interviewees were notified

about their biomarker results by mail. Figure 1 describes the timing of data collection in the HRS panel in two timelines. A line with squares depicts the baseline, and a dashed ellipse illustrates the follow-up. Weir (2008) describes how the integration of biomarkers into the HRS validates and adds information to self-reported health, improves modeling of pathways to health, and allows participants to ascertain what they had not previously known about their health.

Figure 2.1: Timing of biomarker measurements in the HRS



County-level correlates of transitions into diabetes were drawn from the Health and Retirement Study Contextual Data Resource (HRS-CDR). The HRS-CDR is a restricted data set that enables researchers to study the effects of place on health and well-being among the HRS respondents (Ailshire et al., 2020). This study uses the Dartmouth Atlas of Health Care and the 1990-2000 U.S. Censuses. The HRS-CDR Dartmouth Atlas of Health Care documents how

medical resources are distributed and used in the United States, and the Decennial Censuses are comprised of demographic and socioeconomic characteristics on the U.S. population and each of the 3,142 U.S. counties. Data to identify health professional shortage areas were drawn from the County Characteristics Data set, which is produced by the Inter-university Consortium for Political and Social Research (ICPSR, 2008). This study was approved by the University of Illinois at Urbana-Champaign's Institutional Review Board (IRB).

2.3.2. Measurements

To analyze the development of diabetes using logistic models for transitions from prediabetes to diabetes and transitions from no diabetes to prediabetes and diabetes, we classified stages of type 2 diabetes. We used both self-reported diagnosed diabetes and glycated hemoglobin in the biomarker data. The HRS indicated self-reported diagnosed diabetes by asking a respondent to report whether a doctor told the respondent that he/she had diabetes or high blood sugar. If the respondent reported no diabetes, we analyzed the Hemoglobin A1c result to identify undiagnosed diabetes, prediabetes, or normal conditions. Consistent with the American Diabetes Association's (2014) guidelines, individuals with an A1c of lower than 5.7 percent are normoglycemic, which implies no diabetes. An A1c range of 5.7 to 6.4 percent identifies individuals with high risk for future diabetes, to whom the term prediabetic is applied. Because a person with diabetes would have an A1c level of 6.5 percent or higher, we defined this person as an individual with diabetes even if he/she had no self-reported diagnosed diabetes.

One of our principal independent variables is the rurality of the county within which the respondent resides. We used the county as the contextual unit of analysis because the county is small enough to reflect local social and economic conditions (McLaughlin et al., 2001) and is

also a relevant geographic unit for the delivery of many social services such as public health. We divided respondents among three levels of rurality based on the Rural-Urban Continuum Codes classified by the U.S. Department of Agriculture's Economic Research Service (USDA ERS 2003)¹⁸. In our study, (1) Large central/fringe areas are counties in metropolitan areas of 1 million population or more (reference category), and (2) Medium/small areas are counties in metropolitan areas with populations of less than 1 million. To ensure large enough sample sizes in each of the diabetes stages, we combined all nonmetropolitan categories together, hereafter referred to as (3) rural areas. Throughout, we refer to metropolitan counties as urban and nonmetropolitan counties as rural.

2.3.3. Descriptive statistics

Of the 25,331 respondents aged over 50 in the United States (except for Alaska and Hawaii) eligible for the 2006-2012 waves of the HRS interview, 25,327 provided their diabetes status. A total of 18,756 respondents participated in blood-based biomarker data collection in the EFTE interview, and 17,763 submitted their HbA1c values at least once over the 2006-2012 waves. Consenting respondents were biomarked once every four years, and we focused on the 7,276 individuals who appeared both in the baseline and in the follow-up because we aimed to investigate transitions into diabetes. The analyses lastly excluded 721 individuals who had missing data about their geographic information, SES, health behaviors, or county-level contextual variables¹⁹. These lead us to include 6,555 respondents.

¹⁸ The USDA-ERS distinguished metropolitan counties by the population size of their metro area, and nonmetropolitan counties by adjacency of the county to a metro area and the urban population size.

¹⁹ The number of missing covariates in the baseline are listed: geographic information (missing=1), educational attainment (missing=11), employment status (missing=2), insurance status (missing=5), physical activity status (missing=7), BMI (missing=61), smoking history (missing=50), doctor visit (missing=1), and county-level contextual factors (missing=584, including missing for the quality of index =579).

Table 2.1: Descriptive statistics along the Urban-rural Continuum (percentages in each category)

	<u>Prediabetes in baseline</u>			<u>No diabetes in baseline</u>		
	<u>(N=873)</u>			<u>(N=3,466)</u>		
	Large (n=442)	Med/ small (n=255)	Rural (n=176)	Large (n=1,740)	Med/ small (n=1,087)	Rural (n=639)
Having diabetes	28.5	33.3	34.7	34.2	33.2	34.4
Individual characteristics						
Age						
51-64	33.7	29.4	30.1	42.1	42.3	39.6
65-74	40.3	37.7	43.2	36.0	34.2	38.7
75+	26.0	32.9	26.7	21.9	23.5	21.8
Female	57.9	60.0	56.8	58.7	58.6	57.8
Race/Ethnicity						
NHW	63.1	73.3	76.7	79.8	83.6	89.4
NHB	25.1	15.3	18.2	11.3	6.8	6.7
Hispanic	7.7	8.2	2.8	6.8	8.2	2.8
NHO	4.1	3.1	2.3	2.1	1.4	1.1
Educational attainment						
Less than HS	18.3	24.3	25.6	13.3	17.2	17.2
High school	31.9	34.5	43.8	30.8	32.8	43.8
Some college	25.3	24.7	15.3	24.5	23.7	20.2
4 years college+	24.4	16.5	15.3	31.4	26.3	18.8
Married	64.3	65.9	72.2	70.9	73.1	76.8
Employed	32.4	32.9	35.8	41.9	39.3	38.5
Has Medicare	64.5	67.8	69.9	56.3	58.0	6.0
HH Income						
<\$25,000	28.1	29.0	29.6	20.4	25.9	26.8
\$25,000 to \$74,999	48.0	47.8	50.6	43.4	45.0	49.6
≥\$75,000	24.0	23.1	19.9	36.2	29.2	23.6
Smoking history						
Never smoked	13.1	13.7	14.8	12.9	12.1	13.6
Former smoker	40.5	39.2	34.7	45.8	44.2	39.1
Current smoker	46.4	47.1	50.6	41.3	43.7	47.3
Overweight	39.1	41.2	37.5	41.1	41.3	41.8
Obese	38.5	37.3	42.1	23.3	25.2	27.4
Physically active	31.0	31.4	22.2	37.5	37.0	29.9
Visit doctor	96.8	96.1	97.2	95.8	92.6	94.1

Table 2.1: Descriptive statistics along the Urban-rural Continuum (percentages in each category) (cont)

County characteristics						
Quality of health care (index)	1.9	1.9	1.9	1.9	1.9	1.9
HPSA	45.7	46.3	10.2	42.5	39.7	19.6
Population loss	12.9	18.4	4.0	12.0	11.5	8.5
Percentage of seniors	12.2	12.9	13.4	12.1	13.4	14.6
Poor county	6.6	14.5	34.1	3.9	12.1	20.3

Abbreviations: NHW = non-Hispanic Whites; NHB = non-Hispanic Blacks; NHO = non-Hispanic other races

Table 2.1 displays descriptive statistics for our sample. The first three columns show 873 respondents with prediabetes at baseline, and the latter three columns show 3,466 respondents with no diabetes at baseline. For the logistic models estimating the transition from prediabetes to diabetes, we excluded (1) 1,544 respondents who had diabetes in the baseline, and (2) 4,138 respondents who were normoglycemic in the baseline. The result is a total sample of 873 individuals. For the logistic models estimating the transition from no diabetes to prediabetes and diabetes, we excluded (1) 3,089 respondents who had prediabetes or diabetes in the baseline and had a total sample of 3,466 respondents.

Among those with prediabetes in the baseline, 28.5 percent of residents in large central/fringe counties went on to develop diabetes. In contrast, about 34.7 percent of residents in rural counties became people with diabetes in the follow-up. Individuals living in rural counties had the lowest levels of education and were most likely to have household incomes less than \$25,000 per year. With respect to health behaviors, rural residents with prediabetes in the baseline were more likely to be current smokers and to be obese than urban residents with prediabetes. In terms of the characteristics of counties themselves, rural counties were the poorest and had the highest percentage of population aged over 65. Urban counties were more

likely to be designated as primary care health professional shortage areas²⁰ than rural counties. The designation might not characterize an actual shortage but might still allow for federal program support for the area (Monnat and Pickett, 2011).

Among those with no diabetes in the baseline, rural residents were the least likely to have completed four-year college degrees, be employed, and have household incomes over \$74,999 per year compared to individuals in urban counties. They were more likely to report smoking currently and less likely to manage their weight than their counterparts. In terms of the characteristics of counties, rural counties had the highest percentages of seniors, lowest index for quality of health care, and highest percentages of households classified as poor.

2.4 The empirical model

To examine the differences in transitions between stages of type 2 diabetes, we estimated the following logistic equation:

$$Y_{ijt} = \alpha + \beta_1 R_{ijt-1} + \beta_2 X_{ijt-1} + \beta_3 \gamma_{j2000} + \beta_4 \delta_{j2003-2005} + \varepsilon_{it}$$

where Y_{ijt} is equal to 1 when individual i living in a county j in the baseline had developed diabetes by the time of the follow-up t . In the model for the transition from prediabetes to diabetes, observations from middle- and older-aged Americans who had diabetes in the baseline were excluded from the sample. We also excluded the individuals who did not have diabetes in the baseline. The dependent variable is equal to 1 when an individual with prediabetes at baseline goes on to develop diabetes at follow-up and is equal to 0 if an individual with prediabetes keeps

²⁰ Health Professional Shortage Area (HPSA) designations are used to identify areas and population groups within the United States that are experiencing a shortage of health professionals. There are three categories of HPSA designation based on the health discipline that is experiencing a shortage: 1) primary medical; 2) dental; and 3) mental health. The primary factor used to determine a HPSA designation is the number of health professionals relative to the population with consideration of high need. See details: <https://data.hrsa.gov/topics/health-workforce/shortage-areas>

having prediabetes at follow-up. We did not observe the transition of any individual from prediabetes in the baseline to the normal stage in the follow-up. For the model of transition from no diabetes to diabetes, we use an ordered multinomial logit, and the dependent variable is equal to 0 when the individual with no diabetes in the baseline remained normoglycemic in the follow-up, 1 when people with no diabetes or prediabetes in the baseline went on to develop prediabetes in the follow-up, and 2 when people without prediabetes or diabetes in the baseline were classified as people with diabetes at follow-up.

To assess the effects of geographic variation on transitions into diabetes, we classified the county of residence into three types (R_{ijt-1}). The Rural-Urban Continuum Codes (RUCCs or Beale Codes) assigned in the baseline were used to categorize counties using a three-point graduated scale with 1 being the most urban and 3 being the most rural, according to the USDA-ERS (2003)²¹. The regressions included controls for the four regions in the U.S. Census, with South as the reference category. Final variable selection was based on the research summarized above and assessments of multicollinearity and model fit²².

The regressions included the individual-level characteristics (X_{ijt-1}) and the county-level contextual variables that might have caused bias in the estimated coefficients. The individual demographic characteristics included the following race/ethnicity groups: non-Hispanic Whites (reference group), non-Hispanic Blacks, non-Hispanic other races, and Hispanics. Other

²¹ We examined Urban Influence Codes (UICs) as well as the RUCCs. The classification into the three categories of the rural/urban continuum changed only trivially, and using the UICs would result in a loss of observations because county FIPS were missing in the data set.

²² To detect multicollinearity, we calculated the variance inflation factors (VIFs) and the condition number in collinearity diagnostics. None of the VIFs were greater than 10, and the condition numbers were between 4.88 and 5.70 for the models we estimated. We used stepwise models to determine whether certain variables or blocks of variables were accounting for the attenuation of other predictors. There was no single variable that affected the attenuation of rural/urban residence in our models. Using the *oglm* stepwise selection procedure, we determined that heterogeneity was not an issue (Mood, 2010; Kuha and Mills, 2018). Details are available from the authors by request.

demographic factors included gender (male = reference group) and dummy variables for age groups: 51-64 (reference category), 65-74, and 75 and over.

To examine the extent to which the urban-rural differences are explained by differences in individual-level characteristics and whether the socioeconomic status and health behaviors of individuals explain differences in transitions into diabetes from prediabetes or from no diabetes between residents in large central/fringe counties and those in medium/small counties or rural counties, we included several individual-level variables in the baseline *t-1*. Educational attainment was measured using four categories: less than high school graduate, high school graduate (reference group), some college education, and four-year college graduate or above. Dichotomous variables for marital status (married = reference category), employment status (employed = reference category), and Medicare coverage (having Medicare coverage = reference category) were included as controls. Household income was classified into three categories: less than \$25,000, between \$25,000 and \$74,999, and \$75,000 and above (reference category). Health-related behaviors were measured using indicators of smoking, being overweight, being obese, being physically active, and visiting a doctor. Smoking status was classified into three categories: current smokers, former smokers, and people who never smoked (reference category). Being overweight and being obese were dichotomous variables determined by Body Mass Index (BMI). A BMI from 25 to < 30 was classified as overweight, and a BMI of 30 or higher was classified as obese. A respondent was classified as physically active if the person reported exercising once per week or more = 1 (reference category) vs. not = 0. We also controlled for whether the individual visited a doctor within the last two years or not.

Contextual variables represented by the vector γ_{j2000} proxy for county-level structural disadvantages. They included racial/ethnic composition, percent of population aged over 65,

whether the county experienced population loss between 1990 and 2000, whether the county was a poor county, defined as having 20 percent or more of the population living below the poverty line, whether at least one area within the county was designated as a primary care health professional shortage area, and the index for the quality of medical care. To capture any systematic racism such as fewer resources in counties where the minority groups are more likely to live, we included the percent of non-Hispanic Blacks in 2000 and the percent of Hispanics in 2000. We included dichotomous variables indicating whether the county lost population between the 1990 and 2000 Censuses, and whether 20 percent or more of residents were poor.

One important contextual variable that may account in part for residential differences is the availability of health care services ($\delta_{j2003-2005}$). We referred to two sets of nationally accepted preventive service guidelines from the American Diabetes Association (ADA) and Healthy People 2020 to determine the threshold and generate the index of the quality of health care. The ADA provides the recommended frequency of each service, and Healthy People 2020 offered national goals (Table 2.2). We used the Dartmouth Atlas Data set in 2003-2005, which provides the average annual percentage of people with diabetes aged 65-75 who had the HbA1c test, eye exam, and blood lipids test in each U.S. county.

We considered the county to offer high-quality diabetic care if each service variable exceeded the national goal for the services set by Healthy People 2020²³. Our index takes a value between 0 to 3 where we added one point to the index if the average annual percentage of Medicare adults with diabetes aged 65-75 receiving each service is above each threshold. For

²³ Both ADA and Healthy People 2020 do not restrict their targets to middle-aged or older adults, and we classified the highest 25 percent for the HbA1c test, eye exam, and blood lipids test of Medicare adults with diabetes aged 65-75 to set the threshold. The generated thresholds for each test are similar to the national goals, and we decided to use the national guideline to decide our thresholds.

Table 2.2. Nationally Accepted Health Services Guidelines from the American Diabetes Association (ADA) and Healthy People 2020

	ADA	Healthy People 2020
HbA1c Test	Perform the A1c test at least 2 times a year in patients who are meeting treatment goals (and who have stable glycemic control) and quarterly in patients whose therapy has changed or who are not meeting glycemic goals.	≥71.1% of adults aged 18+ with diabetes will have an annual HbA1c test. ≥16.2% of adults aged 18+ with diagnosed diabetes will reduce an A1c value greater than 9%.
Eye Exam	Patients with type 2 diabetes should have an initial eye exam by an ophthalmologist or shortly after the diagnosis. Subsequent eye exams for type 2 diabetes patients should be repeated annually.	≥58.7% of adults aged 18+ with diabetes will have an annual eye exam.
Blood Lipid Test	In adult patients, test for lipid disorders at least annually. In adults with low-risk lipid values (LDL-C<100mg/dl) repeat lipid assessments every 2 years.	≥85.3% of adults aged 18+ with diabetes will have an annual blood lipid test. ≥58.3% of adults aged 18+ with diagnosed diabetes improve lipid control (LDL-C<100mg/dl).

example, Champaign County in Illinois, where the average annual percent of people with diabetes having an A1c test, an eye exam, and blood lipids tests are 84.55 percent (vs. national goal of 84.16 percent), 72.26 percent (vs. national goal of 71.05 percent), and 73.59 percent (vs. national goal of 78.92 percent), respectively, is assigned 2 points. Thus, larger values of the index imply higher quality of care for people with diabetes. A binary indicator for whether any

part of the county was classified as a health care professional shortage area is used to measure health care infrastructure. Access to primary health care may facilitate preventive health interventions. Counties lacking these services may have higher transition rates to diabetes.

2.5. Results

2.5.1. Transitions from prediabetes to diabetes

Table 2.3 presents four models estimating odds of reporting transitions from prediabetes to diabetes. Model 1 displays rurality differences controlling for U.S. Census regions of residence, individual race/ethnicity, dummy variables for age group, and sex. The results demonstrate that as the degree of rurality of the county increases, the probability of experiencing the transition to diabetes increases. People who live in rural areas or medium/small metro counties are more likely to make the transition from prediabetes to diabetes than people who live in large central/fringe metro counties. We also observe that the average resident of rural counties has a greater likelihood of reporting the transition than the average resident in medium/small metro counties, and thus the differences in the transition from prediabetes to diabetes becomes greater with the degree of rurality. There are statistically significant differences in the race/ethnicity coefficients, suggesting that Black Americans with prediabetes at baseline are more likely to go on to develop diabetes than Hispanics and White Americans are.

Individual-level socioeconomic status (SES) indicators including dummies for educational attainment, employment status, marital status, household income, and Medicare coverage are controlled for in Model 2. The addition of variables to control for individual SES does not lead to a remarkable change in magnitude of the county-type variables. Instead, the

Table 2.3. Logistic model to estimate the transition from prediabetes to diabetes

Odd ratios	<u>Pre-diabetes to diabetes (N=873)</u>			
	(1)	(2)	(3)	(4)
Rurality (ref=Large central/fringe counties)				
Medium/small counties	1.41**	1.39**	1.37**	1.37*
	(0.25)	(0.25)	(0.25)	(0.25)
Rural counties	1.50**	1.49**	1.48*	1.47
	(0.31)	(0.31)	(0.31)	(0.35)
Race/Ethnicity (ref=NHW)				
Hispanics	1.31	1.24	1.11	0.98
	(0.39)	(0.39)	(0.36)	(0.32)
NHB	1.25**	1.20*	1.12	1.01
	(0.24)	(0.24)	(0.23)	(0.21)
Female	0.71**	0.70**	0.65**	0.65**
	(0.11)	(0.11)	(0.11)	(0.11)
Individual-level SES				
<i>Educational Attainment (ref=High school)</i>				
Less than HS		0.98	1.11	1.11
		(0.21)	(0.25)	(0.25)
Some college		1.15	1.25	1.26
		(0.23)	(0.26)	(0.26)
4-year college+		0.71	0.73	0.72
		(0.16)	(0.17)	(0.17)
Employed		0.88	0.92	0.94
		(0.16)	(0.17)	(0.18)
Married		1.00	0.99	0.97
		(0.19)	(0.19)	(0.19)
<i>HH Income (ref=HH income > \$74,999)</i>				
HH Income: <\$25,000		1.01	1.05	1.07
		(0.27)	(0.28)	(0.29)
HH Income: \$25,000 to \$74,999		0.95	0.99	1.03
		(0.19)	(0.21)	(0.22)
Has Medicare		1.03	1.07	1.06
		(0.30)	(0.32)	(0.32)
Individual-level Health Behaviors				
Overweight			1.98***	2.00***
			(0.46)	(0.46)

Table 2.3. Logistic model to estimate the transition from prediabetes to diabetes (cont)

Being obese			2.78***	2.80***
			(0.65)	(0.66)
<i>Smoking status (ref=Never smoked)</i>				
Current smoker			1.55*	1.62*
			(0.39)	(0.41)
Former smoker			1.15	1.22
			(0.30)	(0.32)
Physically active			1.24	1.28
			(0.21)	(0.22)
Visited a doctor			0.99	1.06
			(0.42)	(0.45)
County-level contextual variables				
Quality of health care				-0.61*
				(0.16)
HPSA				1.25
				(0.25)
Population loss				0.87
				(0.23)
Percentage of seniors				0.24
				(0.55)
Poor county				1.04
				(0.26)
Model fit statistics				
Pseudo R-squared	0.33	0.38	0.63	0.67
AIC	1,071.40	1,081.76	1,067.14	1,071.13
BIC	1,128.67	1,177.20	1,191.21	1,119.06

Abbreviations: NHW = non-Hispanic Whites; NHB = non-Hispanic Blacks; NHO = non-Hispanic other races; HPSA: Health Professional Shortage area

Notes: All models include controls for dummy variables for age groups: 51-64 (reference category), 65-74, and 75 and over, region of residence: Northeast, Midwest, West, and South (reference category), and non-Hispanic other races. Robust standard errors in parentheses.

***p<0.01, **p<0.05, and *p<0.1

results demonstrate that the introduction of socioeconomic status variables mitigates the statistically significant racial/ethnic disparities in the transition.

Neither does the introduction of individual-level health behaviors in Model 3 result in a dramatic reduction in the magnitudes of the rurality variables. On the other hand, the

introduction of health behavior variables makes the race/ethnicity coefficients lose their significance and attenuates differences in the transition from prediabetes to diabetes among Black Americans on average. The individual-level health conditions that affect the transition from prediabetes to diabetes the most are being overweight and being obese. People who were classified as obese in the baseline were 178 percent more likely to become individuals with diabetes than people who were not obese nor overweight. People who were overweight were 98 percent more likely to become individuals with diabetes than people who were not obese nor overweight.

Model 4 reveals that adding the county-level contextual features to the model further decreases the magnitude of the rurality coefficients. In particular, the finding suggests a reduced difference in the likelihood of the transition between the average residents in a large central/fringe metro county and the average individual living in medium/small metro county, and the medium/small metro county coefficient becomes marginally significant at the 10 percent level. The estimate for rural county is positive and insignificant for the transition from prediabetes to diabetes, in keeping with the expectation that this coefficient captures the partial effects of rurality on the development of diabetes. However, the magnitude of the estimate for rural county does not change much. Controlling for the contextual variables additionally decreases the magnitude of the coefficient for Black Americans. Hence, racial/ethnic differences in the transition are partially explained by differences in individual-level determinants and further explained by differences in county-level characteristics.

Adding county-level contextual variables to our analyses in Model 4 (Table 2.3) caused the effects of rurality to decrease and to become less statistically significant, which is not surprising because these county-level characteristics are related to rurality. The only county-level

contextual variable that approached statistical significance was the quality of health care for diabetes. The result suggests that if the quality of health care for diabetes improves from the worst quality (index value of 0) to the best quality (value of 3), the likelihood of transitioning from prediabetes to diabetes falls by 1.83 percentage points (equal to the coefficient of -0.61×3 units of change in quality).

2.5.2. Transitions from no diabetes to diabetes

Results of the regression analyses for the transition from no diabetes to diabetes are presented in Table 2.4. From all four models, we find little evidence that there are effects of the county types on the direct transition from no diabetes to diabetes. Results of the model regressing the transition on the degree of rurality, U.S. Census regions of residence, race/ethnicity, sex, and age groups (Model 1) reveal statistically significant differences in racial and ethnic coefficients. Compared to a non-Hispanic White American, a Black American has about 197 percent ($OR=2.97$; $[95\%] CI=-1.50$ to 0.67) greater odds of reporting the transition from no diabetes to diabetes. We find that marriage plays an important role to reduce the odds of reporting the transitions to diabetes from no diabetes. There is a significantly lower risk of the development of diabetes for a married person compared to a non-married person.

The introduction of county-level contextual variables in Model 4 leads to an additional change in the magnitude of the racial and ethnic coefficients. In particular, the addition of these variables to the model considerably reduces the magnitude of Black Americans ($OR=2.10$; $[95\%] CI=0.28$ to 1.20) from no diabetes to diabetes. Therefore, racial and ethnic differences in the transition from no diabetes to diabetes are partially explained by differences in the socioeconomic status of minority groups living in certain neighborhoods and further explained by differences in county-level contextual features.

Table 2.4. Multinomial logit model to estimate the transition from no diabetes to prediabetes and diabetes

No diabetes to prediabetes and diabetes (N=3,466)								
	(1)		(2)		(3)		(4)	
Odd ratios	Prediabetes	Diabetes	Prediabetes	Diabetes	Prediabetes	Diabetes	Prediabetes	Diabetes
Rurality (ref=Large central/fringe counties)								
Medium/small counties	0.95	1.07	0.93	1.05	0.93	1.03	0.91	0.96
	(0.09)	(0.18)	(0.09)	(0.19)	(0.10)	(0.18)	(0.10)	(0.18)
Rural counties	0.99	1.19	0.95	1.13	0.95	1.07	0.9	0.94
	(0.12)	(0.25)	(0.11)	(0.24)	(0.11)	(0.23)	(0.11)	(0.22)
Race/Ethnicity (ref=NHW)								
Hispanics	1.31	1.55	1.23	1.26	1.23	1.22	1.26	1.09
	(0.22)	(0.45)	(0.22)	(0.39)	(0.22)	(0.38)	(0.23)	(0.35)
NHB	1.77***	2.97***	1.68***	2.44***	1.62***	2.16***	1.70***	2.10***
	(0.12)	(0.63)	(0.24)	(0.54)	(0.23)	(0.48)	(0.26)	(0.49)
Female	1.11	0.76***	1.04	0.66***	1.12	0.66***	1.12	0.66***
	(0.09)	(0.11)	(0.09)	(0.10)	(0.10)	(0.11)	(0.10)	(0.11)
Individual-level SES								
<i>Educational Attainment (ref=High school)</i>								
Less than HS			0.86	1.25	0.82	1.21	0.82	1.19
			(0.13)	(0.28)	(0.11)	(0.28)	(0.11)	(0.27)
Some college			0.94	1.11	0.94	1.13	0.94	1.14
			(0.11)	(0.22)	(0.10)	(0.23)	(0.11)	(0.23)
4-year college+			0.77**	0.71	0.83	0.79	0.84	0.77
			(0.09)	(0.16)	(0.10)	(0.18)	(0.10)	(0.17)
Employed			0.80**	0.93	0.80**	0.90	0.81**	0.90
			(0.08)	(0.18)	(0.09)	(0.17)	(0.09)	(0.18)
Married			0.93	0.61***	0.95	0.60***	0.96	0.59***
			(0.10)	(0.11)	(0.10)	(0.11)	(0.10)	(0.11)

Table 2.4. Multinomial logit model to estimate the transition from no diabetes to prediabetes and diabetes (cont)

<i>HH Income (ref=HH income>\$74,999)</i>						
HH Income: <\$25,000	1.07 (0.15)	1.01 (0.27)	1.04 (0.15)	0.99 (0.27)	1.04 (0.15)	0.98 (0.28)
HH Income: \$25,000 to \$74,999	0.93 (0.10)	1.27 (0.26)	0.91 (0.10)	1.27 (0.26)	0.90 (0.10)	1.27 (0.26)
Has Medicare	0.96 (0.19)	1.07 (0.38)	0.95 (0.18)	1.00 (0.34)	0.96 (0.19)	1.03 (0.35)
Individual-level Health Behaviors						
Overweight			1.39*** (0.14)	1.97*** (0.43)	1.38*** (0.14)	1.97*** (0.43)
Being obese			1.54*** (0.18)	3.97*** (0.88)	1.52*** (0.18)	3.97*** (0.87)
<i>Smoking status (ref=Never smoked)</i>						
Current smoker			0.64*** (0.09)	0.89 (0.22)	0.64*** (0.09)	0.89 (0.22)
Former smoker			0.79* (0.10)	0.80 (0.20)	0.78 (0.10)	0.80 (0.20)
Physically active			0.87 (0.08)	0.82 (0.14)	0.86 (0.08)	0.82 (0.14)
Visited a doctor			1.04 (0.19)	1.30 (0.45)	1.05 (0.20)	1.28 (0.45)
County-level contextual variables						
Quality of health care (Index)					0.92 (0.12)	1.11 (0.26)
HPSA					0.97 (0.10)	1.04 (0.19)
Population loss					0.86 (0.13)	0.79 (0.21)

Table 2.4. Multinomial logit model to estimate the transition from no diabetes to prediabetes and diabetes (cont)

Percentage of seniors				9.87**	2.80
				(10.66)	(5.89)
Poor county				1.02	1.62*
				(0.17)	(0.42)
Model fit statistics					
Pseudo R-squared	0.12	0.18	0.32	0.34	
AIC	5,096.57	5,097.64	5,046.68	5,057.64	
BIC	5,244.18	5,343.67	5,366.52	5,238.99	

Abbreviations: NHW = non-Hispanic Whites; NHB = non-Hispanic Blacks; NHO = non-Hispanic other races; HPSA: Health Professional Shortage area

Notes: All models include controls for dummy variables for age groups: 51-64 (reference category), 65-74, and 75 and over, region of residence: Northeast, Midwest, West, and South (reference category), and non-Hispanic other races. Robust standard errors in parentheses. ***p<0.01, **p<0.05, and *p<0.10

For the transition from no diabetes to diabetes, one of the county-level proxies that captures structural disadvantage is an important predictor. Specifically, living in a poor county increases the odds of reporting the transition from no diabetes to diabetes. An individual living in a county where 20 percent or more of the residents are poor has about 62 percent (OR=1.62; [95%] CI=-0.035 to 0.99) greater odds of reporting this transition.

2.5. Results

Based on nationally representative data, we investigate the effects of racial, ethnic, and urban-rural differences on transitions from no diabetes and from prediabetes to diabetes. We find that the likelihood of reporting the transition from prediabetes to diabetes increases with the degree of rurality. The significant differences in the transition from prediabetes to diabetes between the residents of large central/fringe metro counties and the residents of rural counties become insignificant after controlling for individual-level characteristics and county-level contextual features. We also find that county-level proxies for structural disadvantage and individual-level correlates attenuate race/ethnicity and rurality disparities in the development of diabetes. Specifically, obesity and overweight are highly correlated with the transition to diabetes, and these two health outcomes are also related to race, ethnicity, and rural status (Zhang et al., 2009; Cohen et al., 2016).

In addition, racial/ethnic differences in transitions into diabetes are partially explained by differences in individual-level factors and further explained by differences in disadvantaged conditions associated with rural counties. As also found in de Oliveira et al. (2020) and Dadgari et al. (2015), we find that marriage is protective against the development of diabetes.

Despite the strengths of this study, there are limitations. First, we explore differences in transitions into diabetes among middle-aged and old-aged Americans living in three types of counties – large central/fringe metro counties, medium/small metro counties, and rural counties. Combining all of the non-metropolitan counties into the category of rural counties may ignore differences in transitions across different types of rural counties. Second, estimating our model requires us to start with a sample of individuals who do not have diabetes in the baseline, and as a consequence, we may underestimate racial, ethnic, and urban-rural differences in transitions into diabetes. For example, a higher proportion of rural residents have already developed diabetes by the baseline. By excluding them from the estimation, we are missing transitions into diabetes earlier in the life cycle, biasing the results. Third, residents of rural counties, Black Americans, and Hispanics were less likely to complete interviews and provide biomarker data than non-Hispanic White Americans and residents of urban counties. This nonrandom selection is also likely to result in an underestimate of racial, ethnic, and urban-rural differences in transitions into diabetes. Fourth, our analysis does not capture the behavior of rural residents who travel outside their community to seek health care when the availability and quality of local health care are limited (Cummings et al., 2013). Indeed, this may explain why we find that indicators for healthcare infrastructure have little effect on transitions to diabetes. Future research should use spatial mapping and analysis techniques to examine differences in transitions to diabetes among rural individuals who travel versus do not travel to metropolitan counties to obtain medical care.

Taken together, the findings remind us that residents of rural counties have the greatest likelihood of reporting the transition from prediabetes to diabetes and show that county-level contextual variables are also important predictors of transitions into diabetes. Therefore, future

public health efforts designed to reduce racial, ethnic, and urban-rural differences in transitions between stages of type 2 diabetes should examine detailed urban-rural classifications while controlling for individual characteristics, health behaviors, and county-level contextual correlates.

2.6. Chapter 2 references

- American Diabetes Association. 2014. "Diagnosis and Classification of Diabetes Mellitus." *Diabetes Care* 37, supplement: S81-S90. <https://doi.org/10.2337/dc14-S081>
- American Diabetes Association. 2018. "Economic Costs of Diabetes in the US in 2017." *Diabetes Care* 41(5):917-928. <https://doi.org/10.2337/dci18-0007>
- Ailshire, Jennifer, Sarah Mawhorter, and Eun Young Choi. 2020. "Contextual Data Resource (CDR): US Decennial Census and American Community Survey Data, 1990-2018" Version 2.0. Los Angeles, CA: USC/UCLA Center on Biodemography and Population Health.
- Basu, Jayasree, Bernard Friedman, and Helen Burstin. 2004. "Managed Care and Preventable Hospitalization Among Medicaid Adults." *Health Services Research* 39(3):489-510. <https://doi.org/10.1111/j.1475-6773.2004.00241.x>
- Boardman, Jason. 2004. "Stress and Physical Health: The Role of Neighborhoods as Mediating and Moderating Mechanisms." *Social Science & Medicine* 58(12):2473-2483. <https://doi.org/10.1016/j.socscimed.2003.09.029>
- Bolen, Julie, Luann Rhodes, Eve E. Powell-Griner, Shayne D. Bland, and Deborah Holtzman. 2000. "State-Specific Prevalence of Selected Health Behaviors, by Race and Ethnicity-- Behavioral Risk Factor Surveillance System, 1997." In: CDC Surveillance Summaries, March 24, 2000. *MMWR* 2000;49(No. SS-2):1-60.
- Boyle, James, Theodore J. Thompson, Edward W. Gregg, Lawrence E. Barker, and David F. Williamson. 2010. "Projection of the Year 2050 Burden of Diabetes in the US Adult Population: Dynamic Modeling of Incidence, Mortality, and Prediabetes Prevalence." *Population Health Metrics* 8(1):1-12. <https://doi.org/10.1186/1478-7954-8-29>
- Brandle, Michael, Honghong Zhou, Barbara R. K. Smith, Deanna Marriott, Ray Burke, Bahman Tabaei, Morton Brown, and William H. Herman. 2003. "The Direct Medical Cost of Type 2 Diabetes." *Diabetes Care* 26(8):2300-2304. <https://doi.org/10.2337/diacare.26.8.2300>

- Ceballos, Rachel M., Gloria D. Coronado, and Beti Thompson. 2010. "Having a Diagnosis of Diabetes Is Not Associated with General Diabetes Knowledge in Rural Hispanics." *The Journal of Rural Health* 26(4):342-351. <https://doi.org/10.1111/j.1748-0361.2010.00299.x>
- Centers for Disease Control and Prevention. 2018. "Long-term Trends in Diabetes. April 2017." <https://stacks.cdc.gov/view/cdc/46096>
- Centers for Disease Control and Prevention. 2020. *National Diabetes Statistics Report, 2020*. Atlanta, GA: Centers for Disease Control and Prevention, U.S. Department of Health and Human Services.
- Cohen, Steven, Sarah K. Cook, Lauren Kelley, Julia Foutz, and Trisha A. Sando. 2016. "A Closer Look at Rural-Urban Health Disparities: Associations between Obesity and Rurality Vary by Geospatial and Sociodemographic Factors." *Journal of Rural Health* 33(2):167-179. <https://doi.org/10.1111/jrh.12207>
- Colberg, Sheri R, Ronald J. Sigal, Bo Fernhall, Judith Regensteiner, Bryan J. Blissmer, Richard R. Rubin, Lisa Chasan-Taber, Ann L. Albright, and Barry Braun. 2010. "Exercise and Type 2 Diabetes: The American College of Sports Medicine and the American Diabetes Association: Joint Position Statement." *Diabetes Care* 33(12): e147-e167.
- Crimmins, Eileen, Jessica Faul, Jung Ki Kim, Heidi Guyer, Kenneth Langa, Mary Beth Ofstedal, Amanda Sonnega, Robert Wallace, and David Weir. 2013. *Documentation of Biomarkers in the 2006 and 2008 Health and Retirement Study*. Ann Arbor, MI: Institute for Social Research.
- Cummings, Janet, Hefei Wen, Michelle Ko, and Benjamin Druss. 2013. "Geography and the Medicaid Mental Health Care Infrastructure: Implications for Health Care Reform." *JAMA Psychiatry* 70(10):1084-1090. doi:10.1001/jamapsychiatry.2013.377
- Dadgari, Atena, Nahid Mazloom, Mohammad Reza Heidari Firouz Abadi, and Imane Bagheri. 2015. "The Relationship between Marital Satisfaction and Compatibility with Type 2 Diabetes." *Iranian Journal of Psychiatry and Behavioral Sciences*, 9(4):e3105. doi: 10.17795/ijpbs-3105
- de Oliveira, Camila, Luciane Viater Tureck, Danilo Alvares, Chunyu Liu, Andrea R.V. R. Horimoto, Mercedes Balcells, et al. 2020. "Relationship between Marital Status and Incidence of Type 2 Diabetes Mellitus in a Brazilian Rural Population: The Baependi Heart Study." *PLoS ONE* 15(8):e0236869. <https://doi.org/10.1371/journal.pone.0236869>
- Dubowitz, Tamara, S.V. Subramanian, Dolores Acevedo-Garcia, Theresa L Osypuk, and Karen E. Peterson. 2008. "Individual and Neighborhood Differences in Diet Among Low-Income Foreign and US-Born Women." *Women's Health Issues* 18(3):181-190. <https://doi.org/10.1016/j.whi.2007.11.001>

- Duncan, Craig, Kelvyn Jones, and Graham Moon. 1996. "Health-Related Behaviour in Context: A Multilevel Modelling Approach." *Social Science & Medicine* 42(6): 817-830. [https://doi.org/10.1016/0277-9536\(95\)00181-6](https://doi.org/10.1016/0277-9536(95)00181-6)
- Eberhardt, Mark, Deborah Ingram, and Diane Makuc. 2001. *Urban and Rural Health Chartbook: Health, United States, 2001*. Hyattsville, MD: National Center for Health Statistics.
- Falco, Gemma, Piero Stanley Pirro, Elena Castellano, Maura Anfossi, Giorgio Borretta, and Laura Gianotti. 2015. "The Relationship between Stress and Diabetes Mellitus." *Journal of Neurology and Psychology* 3(1):1-7. DOI: 10.13188/2332-3469.1000018
- Golden, Sherita Hill, Arleen Brown, Jane A. Cauley, Marshall H. Chin, Tiffany L. Gary-Webb, Catherine Kim, Julie Ann Sosa, Anne E. Summer, and Blair Anton. 2012. "Health Disparities in Endocrine Disorders: Biological, Clinical, and Nonclinical Factors—An Endocrine Society Scientific Statement." *The Journal of Clinical Endocrinology & Metabolism* 97(9):E1579-E1639. <https://doi.org/10.1210/jc.2012-2043>
- Goonesekera, Sunali, May Yang, Susan A. Hall, Shona C. Fang, Rebecca S. Piccolo, and John B. McKinlay. 2015. "Racial Ethnic Differences in Type 2 Diabetes Treatment Patterns and Glycaemic Control in the Boston Area Community Health Survey." *BMJ Open* 12;5(5):e007375. <http://dx.doi.org/10.1136/bmjopen-2014-007375>
- Inter-university Consortium for Political and Social Research (ICPSR). 2008. *County Characteristics, 2000-2007* [United States]. Ann Arbor, MI: Inter-university Consortium for Political and Social Research. <https://doi.org/10.3886/ICPSR20660.v2>
- James, Cara, Ramal Moonesinghe, Shondelle M. Wilson-Frederick, Jeffrey Hall, Ana Penman-Aguilar, and Karen Bouye. 2017. "Racial/Ethnic Health Disparities Among Rural Adults—United States, 2012–2015." *MMWR Surveillance Summaries* 66(SS-23):1-9. <http://dx.doi.org/10.15585/mmwr.ss6623a1>
- Juster, F. Thomas, and Richard Suzman. 1995. "An Overview of the Health and Retirement Study." *Journal of Human Resources* 30(supplement):S7-S56. <https://doi.org/10.2307/146277>
- Kuha, Jouni and Colin Mills. 2018. "On Group Comparisons with Logistic Regression Models." *Sociological Methods & Research* 49(2):498-525. doi:10.1177/0049124117747306
- Macintyre, Sally, Anne Ellaway, and Steven Cummins. 2002. "Place Effects on Health: How Can We Conceptualise, Operationalize and Measure Them?" *Social Science & Medicine* 55(1): 125-139. [https://doi.org/10.1016/S0277-9536\(01\)00214-3](https://doi.org/10.1016/S0277-9536(01)00214-3)

- Mainous, Arch G., Dana E. King, David R. Garr, and William S. Pearson. 2004. "Race, Rural Residence, and Control of Diabetes and Hypertension." *Annals of Family Medicine* 2(6):563-568. <https://doi.org/10.1370/afm.119>
- McLaughlin, Diane, Shannon Stokes, and Atsuko Nonoyama. 2001. "Residence and Income Inequality: Effects on Mortality Among U.S. Counties." *Rural Sociology* 66(4):579-598. <https://doi.org/10.1111/j.1549-0831.2001.tb00085.x>
- Mechanick, Jeffrey, Alan J. Garber, George Grunberger, Yehuda Handelsman, and W. Timothy Garvey. 2018. "Dysglycemia-based Chronic Disease: An American Association of Clinical Endocrinologists Position Statement." *Endocrine Practice* 24(11):995-1011. <https://doi.org/10.4158/PS-2018-0139>
- Mobasser, Majid, Masoud Shirmohammadi, Tarlan Amiri, Nafiseh Vahed, Hossein Hosseini Fard, and Morteza Ghojzadeh. 2020. "Prevalence and Incidence of Type 1 Diabetes in the World: A Systematic Review and Meta-analysis." *Health Promotion Perspectives* 10(2):98-115
- Monnat, Shannon, and Camille Beeler Pickett. 2011. "Rural/Urban Differences in Self-rated Health: Examining the Roles of County Size and Metropolitan Adjacency." *Health & Place* 17(1):311-319. <https://doi.org/10.1016/j.healthplace.2010.11.008>
- Mood, Carina. 2010. "Logistic Regression: Why We Cannot Do What We Think We Can Do, and What We Can Do About It." *European Sociological Review* 26(1):67-82. <https://doi.org/10.1093/esr/jcp006>
- Pearce, Michelle J., Katherine Pereira, and Ellen Davis. 2013. "The Psychological Impact of diabetes: A Practical Guide for the Nurse Practitioner." *Journal of the American Association of Nurse Practitioners* 25(11):578-583. doi: 10.1002/2327-6924.12035
- Robert, Stephanie. 1998. "Community-Level Socioeconomic Effects on Adult Health." *Journal of Health and Social Behavior* 39(1):18-37. <https://www.jstor.org/stable/2676387>
- Rosenblatt, Roger, and L. Gary Hart. 2000. "Physicians and Rural America." *The Western Journal of Medicine* vol. 173(5):348-351. doi: 10.1136/ewjm.173.5.348
- Spanakis, Elias K., and Sherita Hill Golden. 2013. "Race/Ethnic Difference in Diabetes and Diabetic Complications." *Current Diabetes Reports* 13(6):814-823. <https://doi.org/10.1007/s11892-013-0421-9>
- Subramanian, S. V., Ichiro Kawachi, and Bruce Kennedy. 2001. "Does the State You Live in Make a Difference? Multilevel Analysis of Self-Rated Health in the US." *Social Science & Medicine* 53(1):9-19. [https://doi.org/10.1016/S0277-9536\(00\)00309-9](https://doi.org/10.1016/S0277-9536(00)00309-9)

- The US Burden of Disease Collaborators. 2018. "The State of US Health, 1990-2016: Burden of Diseases, Injuries, and Risk Factors Among US States." *JAMA* 319(14):1444-1472. doi:10.1001/jama.2018.0158
- U.S. Census Bureau. 2015. "New Census Bureau Report Analyzes US Population Projections." Available at: <https://www.prnewswire.com/news-releases/new-census-bureau-report-analyzes-us-population-projections-300044527.html> Accessed September 24, 2021.
- U.S. Department of Agriculture, Economic Research Service (ERS). 2003. *Measuring Rurality: Rural-Urban Continuum Codes*. Available at: <https://www.ers.usda.gov/data-products/rural-urban-continuum-codes/>. Accessed December 11, 2020.
- Vogelsang, Eric. 2016. "Older Adult Social Participation and its Relationship with Health: Rural-Urban Differences." *Health & Place* 42:111-119. <https://doi.org/10.1016/j.healthplace.2016.09.010>
- Waidmann, Timothy, and Shruti Rajan. 2000. "Race and Ethnic Disparities in Health Care Access and Utilization: An Examination of State Variation." *Medical Care Research and Review* 57(1-supplement):55-84. <https://doi.org/10.1177/1077558700057001S04>
- Walker, Rebekah Emma Garacci, Anna Palatnik, Mukoso N. Ozieh, and Leonard E. Egede. 2020. "The Longitudinal Influence of Social Determinants of Health on Glycemic Control in Elderly Adults with Diabetes." *Diabetes Care* 43(4):759-766. <https://doi.org/10.2337/dc19-1586>
- Weir, David. 2008. "Elastic powers: The integration of biomarkers into the Health and Retirement Study." *Biosocial Surveys*. National Academies Press (US).
- Zhang, Qi, Youfa Wang, and Elbert Huang. 2009. "Changes in Racial/Ethnic Disparities in the Prevalence of Type 2 Diabetes by Obesity Level Among US Adults." *Ethnicity & Health* 14(5):439-457. <https://doi.org/10.1080/135578508026991552>
- Zhu, Yeyi, and Cuilin Zhang. 2016. "Prevalence of Gestational Diabetes and Risk of Progression to Type 2 Diabetes: A Global Perspective." *Current Diabetes Reports* 16(7):1-11. <https://doi.org/10.1007/s11892-015-0699-x>

CHAPTER 3

DO FOREIGN-EDUCATED NURSES DISPLACE NATIVE-EDUCATED NURSES?

with Mary Arends-Kuenning

3.1. Introduction

The U.S. population is aging rapidly as the baby boomers reach retirement age. While the number of U.S. citizens aged 65 and over is projected to almost double from 43.1 million in 2012 to 83.7 million by 2050, the working population is expected to decline (Ortman et al., 2014). With this aging population, the prevalence of chronic health problems is expected to increase, and these demographic trends will result in a high demand for nursing care. Therefore, labor-market strategies that can satisfy health care demands in the long run are needed.

Hiring foreign-educated nurses is one of the strategies that has been actively used to balance demand and supply of health care (Aiken et al., 2003; Kline, 2003; Lafer, 2005; Tsitouras and Lopez, 2009). The U.S. Congress and Executive branch have used immigration policy to increase the nursing workforce. Policies that vary from year-to-year lead to wide swings in the waiting time for immigrant nurses to obtain work visas. For example, the U.S. government earmarked 50,000 special category permanent employment-based visas for nurses in May 2005 (Arends-Kuenning, 2006), compared to a total of 140,000 permanent employment-based immigrant visas (EB-3) available in 2005 for all skilled workers. In contrast, the Trump administration has been restricting immigration, including through decreases in the number of visas issued (National Foundation for American Policy, 2019), historically high denial rates for H-1B visas (Anderson, 2019), and slow-walking adjudication processes (Hobson, 2019). In

particular, the administration mandated that all applicants for employment-based permanent residency must undergo an in-person interview (Pierce, 2019).

The Trump administration argues that immigration has produced lower wages and higher unemployment for U.S. citizens (Borjas, 2016), and a strategy of recruiting foreign workers is controversial in nursing markets. Several papers have raised questions about the effects of foreign nurse recruitment on quality of patient care in the U.S., the quality of care in the sending countries, and labor market opportunities for native workers (Glaessel-Brown, 1998; Trucios-Haynes, 2002; Martineau et al., 2002; Brush et al., 2004; Lovell, 2006; Aiken et al., 2017). The goals of this paper are to evaluate whether the inflow of foreign-educated nurses in the U.S. nursing market displaces native-educated nurses and to answer whether the short-run supply of native-educated nurses is different from the long-run supply.

An emerging literature contributes to the debate about the relationship between recruiting foreign-educated registered nurses²⁴ and the labor market outcomes of domestic nurses (Schumacher, 2011; Kaestner and Kaushal, 2012; Cortes and Pan, 2014 and 2015). Papers that define immigrant nurses as foreign trained (Schumacher, 2011; Kaestner and Kaushal, 2012) conclude that the effects of immigrant nurses on native nurses' wages are often statistically insignificant, and when they are statistically significant, they are small. Kaestner and Kaushal (2012) find modest displacement effects of foreign-trained nurses on U.S.-trained nurses, but warn that their results might be biased due to weak instruments. Cortes and Pan (2014) define immigrant nurses as foreign born, and they find large displacement effects on native-born nurses, although no significant impacts on wages. The size of the displacement effects of immigrant nurses on native nurses remains as unsettled question. Moreover, studies exploring the effect of

²⁴ Throughout the paper, the use of the term “nurses” refers to registered nurses.

immigrant nurses on the U.S. nursing labor market have not focused on the change covering recent periods. Nursing labor markets may have changed because older nurses postponed their retirement and retired nurses re-entered the labor market in response to the Great Recession (Buerhaus et al., 2009). In addition, the U.S. nursing workforce faces increased demand for health services with the passage of the Patient Protection and Affordable Care Act (ACA) in 2010 (Buerhaus et al., 2017; Dillender et al., 2019). In other words, the labor impacts of foreign-educated nurses on native-educated nurses may have changed since 2008, validating an exploration of the question with more recent data.

To observe whether there is any particular movement in the supply of native-educated nurses caused by an increase in foreign-educated nurses, we use two datasets, the 1980-2000 U.S. Censuses and the American Community Survey (ACS) multi-year aggregates, and several empirical strategies. The sample includes employed registered nurses ages 25 to 70. Because a registered nurse requires greater skill and has experienced greater employment growth as well as lower unemployment than other major nursing occupations (Benson, 2012), this paper considers only registered nurses. In addition, focusing on one specific occupation alleviates empirical problems. Previous studies examining the effects of immigration on a broader range of native workers did not make the adjustment for demand shifts that may confound the relationship between immigration and wages or employment (Kaestner and Kaushal, 2012). Variables that affect the demand for nurses such as the number of physicians or demographic composition in an area are controlled for in this study. Therefore, restricting the sample to registered nurses improves the internal validity of the analysis.

For the empirical strategy, we use an approach based upon Cortes and Pan (2014)'s linear model with OLS and IV methods. There are two main differences between our approach and

Cortes and Pan's approach. First, we focus on foreign-educated nurses, not foreign-born nurses. Second, our instrumental variable for the inflows of foreign-educated nurses per 1,000 population in a commuting zone is the interaction term between the historical distribution by source country of skilled immigrants excluding nurses across areas in the U.S. and the national inflow of new foreign nurses educated abroad. This past settlement instrument helps identify exogenous labor supply shocks. We find no evidence of displacement effects in either the OLS or the IV approach. Instead, we find evidence of complimentary effects in the IV approach – for every foreign-educated nurse that migrates to a commuting zone to work, between 1 and 2 native-educated nurses are additionally employed in the area. The difference between our results and those of Cortes and Pan (2014) is primarily driven by our defining immigrant nurses as foreign educated instead of as foreign born. Because a single instrument may conflate the short- and long-run responses to newly arrived foreign-educated nurses, we implement a multiple instrumentation procedure introduced by Jaeger, Ruist, and Stuhler (2018). Foreign-educated nurses' inflows and their lagged inflows into a commuting zone are instrumented by the past settlement instrument and its lag. We fail to find any statistically significant displacement effects in either the short or the long run. The double IV estimates are smaller than their corresponding conventional IV estimates, and this is consistent across most of the different age and education groups. In terms of native wages, the estimated impact of the lagged foreign-educated nurses is positive and significant. Our results highlight the importance of longer-term adjustment processes and of the sensitivity of results to the choice of instrumental variable.

This study contributes to the literature about the effects of foreign-educated nurses on the US labor market in two ways. First, we approximate local U.S. labor markets by a different geographic unit compared to existing studies that conduct the analysis across metropolitan

statistical areas (MSAs). MSAs cover areas of the United States with major urban population centers and are adjusted periodically to reflect the growth of cities. In contrast, commuting zones (CZs) are a time-consistent measure of local labor markets and cover the entire United States. A commuting zone was excluded if we observed either zero native or zero foreign-educated registered nurses²⁵. The resulting 140 CZs encompass both metro and nonmetro areas and thus provide a better understanding of the nursing labor market in the U.S. than is available using MSAs as the unit of analysis. Table B.1 presents a list of 140 CZs used in this analysis. Second, it is the first study that examines the causal relationship between an increase in foreign-educated nurses and the employment of native-educated nurses with the double instruments. Little evidence of displacement effects observed when using a multiple instrumentation procedure implies that the risks of recruiting foreign-educated nurses are overstated. Therefore, the findings of this paper provide policy implications and stimulate several follow-up research questions.

3.2. Related research

A large body of literature exists on the overall performance of immigrants in the U.S. economy, whereas recently economists focused attention on the labor market for nurses. To provide general trends in the reliance of the U.S. health care system on foreign-born health professionals and to suggest how policymakers and health care practitioners deal with the issue of nursing migration, Arends-Kuenning (2006) examines the relative earnings of registered nurses, licensed practical nurses, and nursing aides over the time period 1990 to 2000. She shows

²⁵ Table C.1 shows the distribution of the number of native and of foreign-educated nurses across commuting zones. Each commuting zone in the sample has at least one native nurse, whereas 217 commuting zone had no foreign-educated nurses sampled in 2015. When we exclude commuting zones that have less than either five U.S.- or five foreign-educated nurses from 1990 to 2015, we have 133 commuting zones as the final areas. Likewise, 120 commuting zones are left if we only include commuting zones with more than both ten native and ten foreign-educated nurses.

the wage differentials by using simple mean comparisons and finds some evidence that foreign-born health professionals have higher annual and hourly wages than their counterparts.

Schumacher (2011) employs econometric analysis and investigates whether immigrant nurses depress the wages of native nurses. He finds a lower wage for non-citizen nurses born outside of the United States by using the Current Population Survey (CPS), which contains information about the country of birth and citizenship status. However, his estimates of the returns to foreign education may be because some immigrant nurses may have obtained training in the States after they migrated. To address this issue, he runs additional regressions with data from the National Sample Survey of Registered Nurses (NSSRN), which identifies nurses who received their training outside the United States. The results indicate a significant wage disadvantage for newly arrived foreign-educated nurses. However, his research on the effects of immigration on the wages of native nurses did not address the non-random location decisions of foreign nurses. Kaestner and Kaushal (2012) apply an instrumental variables approach with the same NSSRN dataset. They adapt the methods of Altonji and Card (1991), using the lagged numbers of the foreign-educated nurses to instrument for the current supply of nurses, concluding that there is no statistically significant relationship between wages and an increase in supply induced by foreign-educated nurses.

Cortes and Pan (2014) also attempt to alleviate the endogeneity issue by an alternative instrument in the nursing market. They exploit the U.S. Censuses for 1980-2000 and the American Community Survey three-year aggregates for 2010 to estimate the displacement effects of an increase in the supply of foreign-born nurses per capita on the employment of native nurses. For the instrument, they interact the historical distribution of skilled migrants by country of origin across MSAs with the total number of foreign-born nurses from the country. The

standard shift-share IV interacts the historical distribution with the inflow of newly arrived immigrant nurses, not the stock at one point in time²⁶. Their IV estimates imply that one to two native nurses are displaced by each foreign-born nurse and show little evidence that wages fall due to foreign-born nurses. To check whether the quality of co-worker interactions in a workplace is an important factor affecting the labor supply decision, the authors use data from the 2006 and 2010 Survey of Registered Nurses conducted by California Board of Registered Nursing. They regress two dependent variables, support from other nurses and teamwork with other nurses on the share of foreign-educated nurses at the county level in California. They estimate regressions separately for native and foreign-educated nurses. They cannot calculate an instrument for the share of foreign nurses because the sample size is too small. They find that the share of foreign-educated nurses affects domestic nurses and foreign nurses in opposite directions and suggest evidence that a falling-off in the quality of working conditions partially drives the displacement. Cortes and Pan (2015) analyze the quality differentials of nurses educated abroad and measure wages as a proxy of skill. They conclude that foreign-educated nurses are on average more productive than native nurses, and Filipino nurses gain a large and highly statistically significant wage premium.

Combining the national inflow of immigrants with their past geographic distribution is a well-known instrument to identify exogenous labor shocks in the literature on immigration. However, relying heavily on a single instrument may cause inconsistent results because of the compound effect generated by the short- and long-run responses to immigrant arrivals (Jaeger et al., 2018). General equilibrium adjustments and the high correlation between the country of

²⁶ The Cortes and Pan (2014) article is ambiguous about the instrument that they used. On page 165, the description implies that they used a shift-share instrument. However, on page 168, the description given in the text and Equation (3.2) indicate that the authors used the stock of foreign nurses from a specific country in year t , rather than the flow of nurses that entered the U.S. from time $t-1$ to t .

origin composition and immigrant's settlement patterns are key factors that result in this problem. They argue that multiple instruments ought to be used to avoid conflating a fall in wages when new immigrants enter and a positive movement toward equilibrium. Instrumenting two endogenous variables with the past settlement instrument and its lagged instrument, they address the endogeneity of current and past immigrant inflows to current and past labor demand shocks. The base period to construct the two instruments is 1970 when the country of origin of immigrants was less concentrated and showed more variation than in subsequent years. Also, the authors use U.S. Census data starting from 1980 to avoid a mechanical relationship between the lagged independent variable and the instrument because the instrument is a function of past values of the lagged independent variable. For example, the instrument is correlated with the lagged independent variable if it is determined by the historical distribution of the immigrant share in 1980. They find a negative effect and a positive effect of immigration on natives' wage in the short- and long-run, respectively. Their strategy can be used in the study of the U.S. nursing labor market, and hence this paper separates the initial response and the long-period effects of an increase in foreign-educated nurses on native nurses by adding a lagged past settlement instrument to the regression.

3.3. Data and descriptive statistics

The 1980-2000 U.S. Censuses and the American Community Survey three-year aggregates for both 2007 (2005-2007) and 2010 (2008-2010) and five-year aggregates for 2015 (2011-2015) are used as the data sources. All were obtained through IPUMS-USA (Ruggles et al. 2017). The decennial Census samples and the ACS sample for 2015 include 5 percent of the U.S. population, and the ACS samples for 2007 and 2010 include 3 percent of the population. They

contain information on personal and demographic characteristics, the country of birth, earnings, and employment. However, their major limitation is the absence of data on the country of education²⁷. We assume that a foreign-born registered nurse was educated in the U.S. if the nurse was younger than 25 years old when he or she first arrived to live in the United States. The age of 25 was chosen to allow for completion of a Bachelor's degree along with two years of hospital nursing experience required to obtain a visa to work as a registered nurse in the U.S. According to a person's specific year to enter the States and the age of this person in the relevant period surveyed, we identify U.S.- and foreign-educated registered nurses²⁸. In this process, we will miscode those who educated to become a nurse after arriving in the U.S. and those who completed training in their country of origin and arrived in the U.S. before age 25. Nurses born in the United States are assumed to have completed their education in the States.

The samples that represent the population include all individuals not residing in group quarters, while the employment samples are restricted to employed workers aged 25 to 70 who reported their occupation to be that of a registered nurse. The population and employment samples are weighted by the Census personal weights. The sample only includes nurses who worked a positive number of hours and weeks in the previous year and received a strictly positive hourly wage. The hourly wage is calculated as annual earnings divided by the product of usual hours worked per week and weeks worked last year. The wage samples are weighted by the product of the Census personal weight and the number of weeks worked in the previous year. The "weeks worked last year" variable for 2010 and 2015 is a categorical variable, intervals of

²⁷ The National Sample Survey of Registered Nurses (NSSRN) includes information about whether the nurse was educated in the U.S. or another country, and previous literature used this data set. We cannot use it because it was discontinued in 2008. It is also a small data set relative to the Census and ACS.

²⁸ For example, a nurse whose age is 32 in 2000 and entered into the States to live in 1992 is assumed to be a U.S.-educated nurse because the person's age was calculated as 24 years old when he/she first arrived in the States.

weeks worked (such as 14-26 weeks or 50-52 weeks). In contrast, the “weeks worked last year” for 1980-2007 is the precise number of weeks worked. Thus, each variable for 2010 and 2015 is converted to a continuous variable by using the 50th percentile of each interval in 2007. The annual earnings variable includes wages, salaries, commissions, cash bonuses, or tips from all jobs, before tax deductions. Wages are deflated so that they are in terms of the 1999 wages using the BLS CPI-U. Wages that were calculated to be below the minimum wage are set to the minimum wage, and wages calculated to be above top-coded yearly wages are set to dollars²⁹ adjusted by the Consumer Price Index Research Series Using Current Methods (CPI-U-RS). We trim hourly wage outliers.

Previous studies have often used Metropolitan Statistical Areas (MSAs) as a proxy for US nursing labor markets (e.g., Kaestner and Kaushal, 2012; Cortes and Pan, 2014). MSAs are defined by the U.S. Office for Management and Budget for statistical purposes and adjusted periodically. However, geographic inconsistency is problematic for an analysis of the changes in employment composition over time. In addition, MSAs only cover areas of the States with major urban population centers. Hence, the main geographical unit of this paper is Commuting Zones (CZs), which are time-consistent measures of local labor markets developed by Tolbert and Sizer (1996). Commuting zones are clusters of counties that are characterized by strong within-cluster and weak between-cluster commuting ties. To match the geographic information contained in the IPUMS data to CZs, the crosswalk developed by David and Dorn (2013) is applied. We multiply the person weights with an adjustment factor that accounts for the fraction of a Census Public Use Micro Area (PUMA) that maps to a given commuting zone.

²⁹ For example, hourly wages which are above \$184.55 that are censored at the top are set to \$184.55 for 2015 (Economic Policy Institute, 2018).

Table 3.1: Demographic and labor supply characteristics of U.S.-educated and foreign-educated nurses, 1980-2015

	U.S.-educated RNs					
	1980	1990	2000	2007	2010	2015
Age (mean)	39.25	40.01	42.90	44.71	44.99	44.60
Gender						
Female	95.91	94.48	92.44	91.29	91.12	90.20
Male	4.09	5.52	7.56	8.71	8.88	9.80
Race						
White	90.14	88.18	85.01	83.38	82.22	81.17
Black	7.97	8.66	8.71	9.34	9.95	10.10
Asian	1.46	1.94	3.64	4.33	5.11	5.41
Native American	0.30	0.39	0.40	0.45	0.40	0.40
Other Races	0.12	0.83	2.25	2.50	2.32	2.93
Hispanics	1.84	2.67	3.26	4.17	5.10	5.96
Marital Status						
Married	67.13	67.53	68.14	65.88	65.40	63.96
Widow/divorce/separated	18.26	18.79	19.56	20.56	20.25	19.09
Single	14.61	13.68	12.30	13.56	14.35	16.95
Educational Attainment						
Less than HS	12.35	0.56	0.21	0.07	0.21	0.19
High School	6.06	15.43	11.27	6.45	6.69	6.00
Associate's degree	48.51	38.38	35.96	37.32	37.40	34.65
Bachelor degree	21.91	32.45	38.75	42.01	42.66	44.83
Graduate degree	11.18	13.19	13.80	14.15	13.04	14.33
Spanish speaking	1.98	2.91	3.62	3.91	4.15	4.50
Part-time	31.00	30.22	26.32	25.13	23.37	21.70
Hourly wage (mean)	16.24	20.65	22.89	24.42	24.47	23.97
Hours worked/wk (mean)	34.67	36.24	37.11	37.46	37.52	37.72
Number of observations	51,739	81,944	100,102	72,886	78,993	137,429
Estimated number	1,034,780	1,638,880	2,002,040	2,429,533	2,633,100	2,748,580

Source: The data is from the U.S. Censuses and the American Community Survey (ACS) multi-year aggregates.

Note: The Census samples and ACS sample for 2015 include 5 percent of the U.S. population while the ACS sample for 2007 and 2010 include 3 percent of the population. The samples are restricted to employed registered nurses aged 25-70 years old. Wages are deflated so that they are in terms of the 1999 wages using the BLS CPI-U.

Table 3.1: Demographic and labor supply characteristics of U.S.-educated and foreign-educated nurses, 1980-2015 (Cont)

	Foreign-educated RNs					
	1980	1990	2000	2007	2010	2015
Age (mean)	41.15	44.00	45.74	46.61	47.12	48.16
Gender						
Female	93.63	93.88	89.48	86.75	86.36	84.33
Male	6.37	6.12	10.52	13.25	13.64	15.67
Race						
White	36.33	27.22	24.02	20.66	19.36	19.50
Black	16.56	18.52	20.65	21.57	20.14	22.84
Asian	45.96	52.30	49.14	54.39	57.19	54.03
Native American	0.16	0.22	0.18	0.06	0.18	0.20
Other Races	0.99	1.74	6.01	3.32	3.14	3.43
Hispanics	8.49	7.25	5.12	5.11	5.34	5.62
Marital Status						
Married	64.27	68.74	71.68	71.53	73.32	72.86
Widow/divorce/separated	14.55	15.63	15.03	16.27	15.32	16.79
Single	21.18	15.63	13.29	12.20	11.36	10.34
Educational Attainment						
Less than HS	17.50	1.42	0.80	0.60	0.58	0.53
High School	5.09	15.30	11.19	6.06	5.15	5.05
Associate's degree	26.12	24.98	22.77	20.45	19.85	18.92
Bachelor degree	27.46	43.57	51.06	57.98	61.15	61.01
Graduate degree	23.84	14.74	14.19	14.91	13.27	14.48
Spanish speaking	6.29	6.15	5.44	5.27	5.11	5.77
Part-time	17.60	16.89	15.14	12.07	12.89	13.77
Hourly wage (mean)	18.36	24.41	27.49	28.16	27.52	27.87
Hours worked/wk (mean)	37.29	38.60	39.04	40.12	39.48	39.10
Number of observations	3,241	4,459	5,460	4,955	5,871	10,332
Estimated number	64,820	89,180	109,200	165,167	195,700	206,640

Source: The data is from the U.S. Censuses and the American Community Survey (ACS) multi-year aggregates.

Note: The Census samples and ACS sample for 2015 include 5 percent of the U.S. population while the ACS sample for 2007 and 2010 include 3 percent of the population. The samples are restricted to employed registered nurses aged 25-70 years old. Wages are deflated so that they are in terms of the 1999 wages using the BLS CPI-U.

Table 3.1 compares the demographic and labor supply characteristics for U.S.- and foreign-educated registered nurses. The foreign-educated RNs are more likely to be older and

male than US-educated nurses. The average age of nurses and the proportion of nurses who are male have been increasing over time, suggesting that the RN workforce is aging and gender role attitudes have liberalized. The foreign educated earn higher hourly wages than their counterparts. One explanation for wage differentials in favor of internationally educated nurses is that the foreign educated have higher levels of education on average than the U.S.-educated nurses. Each year the foreign educated are more likely to have at least a Bachelor's degree than domestic nurses. In 2015, only 45% of the U.S.-educated RNs had a Bachelor's degree, compared to 61% of the foreign-educated RNs. The foreign educated work more than the U.S. educated and are also less likely to work part time, which is defined as working 35 hours or less during a typical week. Thus, the foreign-educated nurses show greater attachment to the workforce, which would tend to result in higher hourly wages.

Table 3.2 presents the top ten source countries for employed foreign-educated registered nurses from 1980 to 2015. These ten countries account for close to 70% of the total number of foreign-educated nurses. The largest country of origin is the Philippines. The dominance of the Philippines as a source country can be explained by its history as a colony of the U.S. and its government policy that encourages overseas employment as a source of funds (Yang and Martinez, 2006; Garchitorena, 2007; McDonald and Valenzuela ,2012). India has been second at 7% from 2000, followed by Canada. Canada was the second largest among the source countries, but the percentage of nurses educated in Canada is steadily falling. This table illustrates that there is no large change in the percentage of total foreign-educated nurses accounted for by the top ten source countries from 1980 to 2015 and the country of origin composition for foreign-educated nurses might not vary much over time.

Table 3.2: Top ten source countries for employed foreign-educated RNs in the U.S., 1980-2015

1980 (N=3,241)			1990 (N = 4,459)			2000 (N = 5,460)		
Birthplace	N	%	Birthplace	N	%	Birthplace	N	%
Philippines	883	27.25	Philippines	1,566	35.11	Philippines	1,767	32.37
Canada	354	10.91	Jamaica	344	7.72	India	398	7.29
India	199	6.15	Canada	298	6.69	Canada	373	6.83
Jamaica	195	6.01	India	270	6.05	Jamaica	355	6.51
Korea	148	4.56	Korea	154	3.46	Nigeria	217	3.97
England	107	3.29	Ireland	115	2.57	Korea	150	2.75
Ireland	101	3.11	England	102	2.29	Guyana/British Guiana	116	2.13
Thailand	74	2.28	Trinidad and Tobago	89	2.00	England	116	2.12
Trinidad and Tobago	64	1.99	Thailand	77	1.73	China	112	2.05
West Germany	59	1.83	Barbados	60	1.34	Trinidad and Tobago	96	1.77
Percentage of total FENs accounted for by top 10 source countries		67.38			68.96			67.79
2007 (N=4,955)			2010 (N=5,871)			2015 (N=10,332)		
Birthplace	N	%	Birthplace	N	%	Birthplace	N	%
Philippines	1,792	36.17	Philippines	2,335	39.77	Philippines	3,910	37.84
India	421	8.49	India	476	8.11	India	797	7.72
Canada	251	5.07	Jamaica	2665	4.53	Nigeria	593	5.64
Nigeria	229	4.63	Nigeria	262	4.46	Jamaica	404	3.91
Jamaica	220	4.44	Canada	229	3.91	Canada	381	3.69
Korea	162	3.26	Korea	191	3.26	Korea	282	2.73
Haiti	142	2.86	China	145	2.47	China	232	2.25
China	122	2.47	Haiti	121	2.06	Haiti	203	1.97
Trinidad and Tobago	74	1.49	Guyana/British Guiana	70	1.20	Ghana	196	1.90
England	70	1.42	Cuba	70	1.19	Cuba	148	1.44
Percentage of total FENs accounted for by top 10 source countries		70.30			70.96			69.09

Source: The data is from the U.S. Censuses and the American Community Survey multi-year aggregates.

3.4. Methodologies and results

To examine the causal effects of immigration of foreign-educated nurses on the employment of domestic nurses in the United States, the analysis closely follows a linear model that Cortes and Pan (2014) used. It basically employs the “cross-area approach” proposed by Card (2001) with the assumptions that areas represent separate labor markets and that employment of a registered nurse is determined by supply and demand factors in those markets. We exclude commuting zones that have zero registered nurses among 741 commuting zones that encompass the 51 U.S. states. Thus, 140 commuting zones are finally used for the analysis. These CZs cover only 19 percent of the nation’s land, but they cover 80 percent of the U.S. workforce³⁰. The empirical specification is as follows:

$$(3.1) \quad \left(\frac{USEN}{Population} \right)_{it} = \alpha + \beta \left(\frac{FEN}{Population} \right)_{it} + \gamma X_{it} + \delta_i + \lambda_t + \varepsilon_{it}$$

where i is the commuting zone, and t is the time period ($t = 1990, 2000, 2007, 2010$, and 2015). For the main dependent and independent variables, the number of full-time employed U.S.- and foreign-educated nurses³¹ per 1,000 population in a commuting zone i in each time period t is focused. To compute full-time employment, we include nurses working at least 35 hours per week and one-half of nurses working less than 35 hours per week.

Similar to Cortes and Pan (2014), we assume that the U.S.-educated nursing labor force is a function of the number of full-time employed foreign-educated nurses per 1,000 population and time-varying commuting zone level characteristics (X_{it}). Time-varying commuting zone level controls consist of proxies for demand and supply determinants. The demand-side factors include

³⁰ 120 commuting zones, excluding areas with either less than ten native nurses or ten foreign-educated nurses from 1980 to 2015, cover 17 percent of the nation’s mainland, but they cover 77 percent of the U.S. workforce (Table B.2).

³¹ For expositional convenience, we refer to the U.S.-educated nurses and foreign-educated nurses as the “USENs” and “FENs”, respectively.

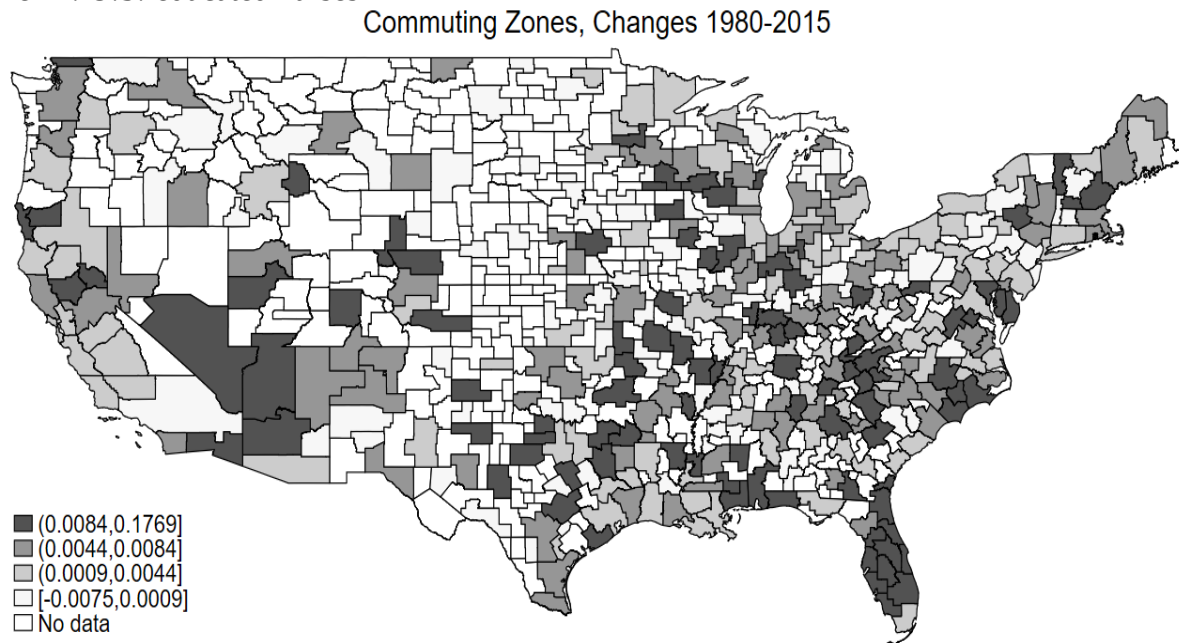
the share of the population aged 65 and over, the log of average hourly wages to describe a commuting zone's income level, the number of physicians in the labor force per 1,000 population, and the number of Spanish speakers at home per 1,000 population. The supply-side factors of the U.S. nursing labor market include the share of the population aged 25 to 29, 30 to 34, 35 to 39, 40 to 44, 45 to 49, 50 to 54, 55 to 59, and 60 to 64 years, the number of females in the labor force who have professional occupations³² excluding registered nurses, the labor force participation of skilled married females, the share of skilled whites in the labor force in the population, the log average hourly wages of employed women who have received some college or more education, excluding registered nurses. Commuting zone (δ_i) and year (λ_t) fixed effects are also included in a specification.

There are two widely recognized concerns with this strategy. First, local markets are not closed, and natives may respond to the immigrant supply shock by moving to other areas. For example, domestic nurses in a commuting zone that received many immigrant nurses may escape the economic conditions by moving to locations with few immigrants. However, a native mobile response to immigration has not been strongly supported by empirical evidence. Card and Lewis (2007) argue that native outmigration in response to an inflow of immigrants appears to be relatively small. Second, immigrants are not randomly distributed across labor markets. Immigrants tend to select locations where previous immigrants have settled because they could lower the cost of job searching and moving to a new place given information networks (Munshi, 2003). Figure 3.1 shows the geography of the change in employed nurses as share of initial population for the 1980-2015 period indicating with darker shades of grey the CZs with larger increases of registered nurses. The change in native nurses as share of initial population is less

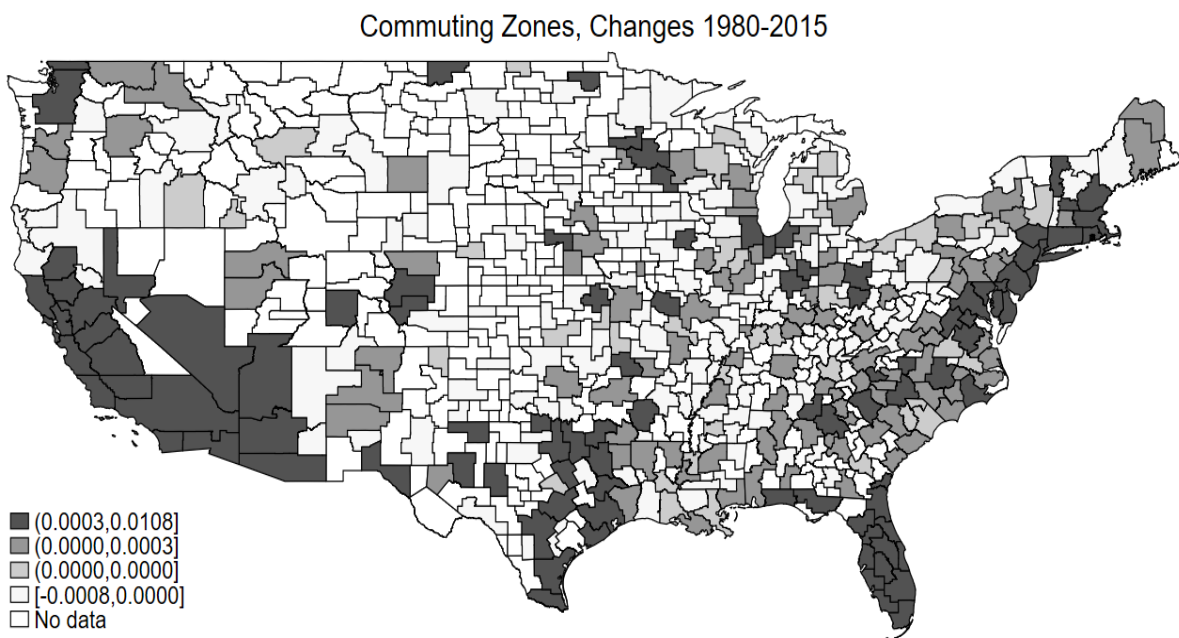
³² Occupations are grouped by BLS, and it defines that most professional occupations require educational preparation (For lists of professional occupation, visit: <https://www.bls.gov/ncs/ocs/ocsm/commoga.htm>).

Figure 3.1. Changes in registered nurses as share of initial population

Panel A: U.S.-educated nurses



Panel B: Foreign-educated nurses



Note: The data is from the U.S. Censuses and the American Community Survey (ACS) multi-year aggregates. The samples are restricted to employed registered nurses aged 25-70 years old, and we plot the maps for US mainland only. These maps show that growth in foreign-educated nurses is not randomly distributed across the United States.

concentrated in specific areas than the foreign educated. Several CZs in California, Texas, and the east coast show the largest values for internationally educated nurses. If the areas where immigrants cluster have done well over time periods, this produces a spurious correlation between immigration and area outcomes. For example, a positive correlation indicates that immigrants choose to live areas where the labor outcomes are relatively good³³. This correlation could obscure the hypothesized effects of the inflow of foreign-educated nurses on their counterparts even though the regression model includes a vector of area fixed effects.

To control for this possible endogeneity bias, the tendency of foreign-educated nurses to settle in an area with a large group of immigrants from the same country of origin is commonly used. The instrument is an interaction between the historical distribution of skilled migrants, excluding registered nurses, across U.S. commuting zones and the national inflow of foreign-educated nurses aged 25 to 70 years old in the labor force. Formally, the instrument for the number of foreign-educated nurses in a commuting zone i and year t is as follows:

$$(3.2) \quad \tilde{m}_{it} = \sum_c \frac{\text{Skilled Immigrants}_{ci,1980}}{\text{Skilled Immigrants}_{c,1980}} * \Delta FEN_{ct,-i}$$

where c refers to the country of birth, i the commuting zone, and t time period ($t = 1990, 2000, 2007, 2010, \text{ and } 2015$). An immigrant is identified based on the country of birth, and a skilled immigrant is defined as an individual who reports having at least an Associate's degree and being in the labor force. The first term in the interaction is the fraction of skilled immigrants from the same country of origin in a commuting zone i at reference date 1980, and the second term is the total number of newly arrived foreign-educated nurses from that country of birth between t and $t-1$ at the national level, net of the contribution of commuting zone i to the national inflow. There

³³ Cadena and Kovak (2016) demonstrated that location choices of low-skilled Mexican-born immigrants in the U.S. are strongly influenced by changes in local labor market conditions.

is a potential concern that a few large areas drive the national import of nurses, and the national inflow of foreign-educated nurses is not orthogonal to the local demand conditions. To reduce this issue, we exclude the contribution of each observation's area to the national inflows in each time period. A justification of this "past settlement instrument" is that location choices in 1980 of skilled immigrants, excluding registered nurses, are exogenous to changes in the demand for foreign-educated nurses over time periods, and that a change in the national inflow of foreign-educated nurses in a given period is not endogenous to local conditions.

Estimates for the first-stage regression of the foreign-educated nurses per 1,000 population in a commuting zone on the instrument (the predicted number of foreign-educated nurses in the commuting zone per 1,000 population) are presented in Table 3.3. The estimates are statistically significant and robust to the inclusion of the controls and fixed effects. The coefficients on the instrument for the 1990-2015 time period (first five columns) and the 2000-2015 time period (last five columns) indicate that as the predicted number of foreign-educated nurses in a commuting zone increases by 10, approximately three to six foreign-educated nurses would enter into the commuting zone. In the whole samples, we do not have a weak instrument problem. Stock and Yogo (2002) report that a first-stage F statistic greater than 8.96 of the Stock and Yogo threshold on 15% tolerable bias level is sufficient to reject the null hypothesis of a weak instrument.

Table 3.4 presents the estimates of the relationship between the importation of foreign-educated nurses and the employment of native-educated nurses in Equation (3.1). Panel A reports the Ordinary Least Squares (OLS) estimates, and Panel B displays the Instrumental Variable (IV) estimates for the displacement regression. All regressions are weighted by the commuting zone's population and standard errors are clustered at the commuting zone level. The

Table 3.3: The first stage regressions

	Dep var: Foreign-educated nurses per 1,000 population									
	I. 1990-2015					II. 2000-2015				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Predicted Foreign-educated RNs per 1,000 population	2.414***	0.521**	0.296**	0.574*	0.241	2.493***	0.369**	0.232*	0.362	0.241
	(0.604)	(0.220)	(0.166)	(0.308)	(0.348)	(0.628)	(0.214)	(0.183)	(0.415)	(0.348)
Controls		X	X	X	X		X	X	X	X
Year FE	X	X	X	X	X	X	X	X	X	X
Area FE			X	X	X			X	X	X
Pre-Great Recession				X					X	
Great Recession and Post-Great Recession					X					X
F-statistics	496.73	32.05	10.21	23.76	2.35	341.64	20.63	10.23	5.43	2.35
Observations	700	700	700	420	280	560	560	560	280	280
R-squared	0.453	0.817	0.943	0.967	0.973	0.410	0.813	0.947	0.977	0.973

Source: The data is from the U.S. Censuses and the American Community Survey multi-year aggregates.

Note: The dependent variable in the first-stage regressions is the number of foreign-educated nurses aged 25-70 as a fraction of a whole population in a commuting zone. The instrument is constructed by using the historical distribution of skilled immigrants (no restrictions on age or labor force status), excluding those who are registered nurses, across commuting zones in 1980 to allocate the national inflow of foreign-educated nurses to each commuting zone. All specifications are weighted by the whole population in the commuting zone. Standard errors in parentheses are clustered at the commuting zone level. The symbol * means significant at 10% level, ** at 5% level and *** at 1% level.

difference between Columns (1)-(5) and (6)-(10) is the exclusion of observations from 1990 (or the 1990 sample)³⁴. The OLS estimates are negative and statistically significant without the commuting zone fixed effects, and the magnitudes with control variables indicate that the inflow of 10 internationally educated nurses in a commuting zone causes the displacement of 5 to 6 native-educated nurses. However, the estimates become positive and statistically insignificant in the preferred specification (Column (3) and (8)) that includes controls and two fixed effects. The OLS estimates are likely to be confounded by demand or supply shocks to the U.S. nursing market, and thus cautious interpretation is needed.

To identify exogenous labor shocks, the instrumental variable (IV) strategy is employed. Panel B reports the 2SLS estimates where we instrument for the number of full-time employed foreign-educated nurses per 1,000 commuting zone population using the past settlement instrument (Equation (3.2)). The IV estimates are negative with the inclusion of controls and year fixed effects. The magnitudes range from -3.018 to -1.902 indicating that for every foreign-educated RN who enters a commuting zone, the number of native-educated nurses decreases by 2-3. We include area fixed effects to circumvent the issue that the baseline geographic sorting of foreign-educated nurses originating in a country c was partly determined by pre-existing regional difference in employment opportunities. The IV estimates are positive and significantly different from zero with the full sample (Column (3) and (8)), and the magnitudes imply that for every foreign-educated nurse that enters a commuting zone, approximately 1-2 native nurses are employed. The results indicate that the instrument is likely to be confounded by unobserved demand and supply shocks, thereby justifying the use of the multiple instrumentation procedures.

³⁴ The reason for estimating with two different samples is to compare the latter results with the outcomes obtained by using a multiple instrumentation procedure.

Table 3.4: Displacement effects of foreign-educated RNs on U.S.-educated RNs

	Dep var: U.S.-educated nurses per 1,000 population									
	I. 1990-2015					II. 2000-2015				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
OLS										
Foreign-educated nurses per 1,000 population	- 1.730*** (0.347)	-0.636** (0.250)	0.215 (0.286)	-0.020 (0.261)	0.605 (0.381)	- 1.816*** (0.339)	-0.466* (0.254)	0.205 (0.263)	-0.035 (0.398)	0.605 (0.381)
Observations	700	700	700	420	280	560	560	560	280	280
R-squared	0.390	0.778	0.932	0.954	0.968	0.390	0.790	0.946	0.975	0.968
2SLS (Single Instrument)										
Foreign-educated nurses per 1,000 population	- 1.988*** (0.465)	- 1.902*** (0.737)	2.182* * (0.995)	0.107 (0.740)	0.970 (1.487)	- 2.223*** (0.476)	-3.018* (1.734)	1.878* (1.355)	0.237 (1.495)	0.970 (1.487)
Observations	700	700	700	420	280	560	560	560	280	280
R-squared	0.383	0.746	0.909	0.954	0.967	0.374	0.664	0.931	0.975	0.967
Controls		X	X	X	X		X	X	X	X
Year FE	X	X	X	X	X	X	X	X	X	X
Area FE			X	X	X			X	X	X
Pre-Great Recession				X					X	
Great Recession and Post-Great Recession					X					X

Source: The data is from the U.S. Censuses and the American Community Survey multi-year aggregates.

Note: The dependent variable is the number of full-time employed U.S.- educated nurses aged 25-70 years per 1,000 population in a commuting zone, and the main independent variable is the number of full-time employed internationally educated nurses aged 25-70 years per 1,000 population in a commuting zone. For the specification, the 2SLS model including controls and fixed effects is applied. All specifications are weighted by the whole population in the commuting zone. Standard errors in parentheses are clustered at the commuting zone level. The symbol * means significant at 10% level, ** at 5% level and *** at 1% level.

Because our results are so different from those of Cortes and Pan (2014), who found large employment displacement effects of foreign-born nurses on native-born nurses, we examine the source of the difference (Results are available from the authors by request). We conclude that the difference is driven by using the definition of foreign educated instead of foreign born for immigrant nurses. We found negative effects of the supply of foreign-born nurses on the supply of US-born nurses using the IV strategy. The difference was not driven by our inclusion of the most recent years of data. We found no difference between the Great Recession and the post-Great Recession period (Table 3.4). We also concluded that the difference in results was not driven by using commuting zones instead of MSAs as the geographic area of analysis. We found a positive effect of foreign-educated nurses on U.S.-educated nurses as a negative effect of foreign-born nurses on U.S.-born nurses when using metropolitan areas as the unit of analysis. The definition of immigrant nurses makes a difference for the results.

We additionally investigate which native nurses are the most affected by foreign-educated nurses. Table 3.5 and Table 3.6 present the effects by different age groups and education levels, respectively. Only the 2SLS model including controls and fixed effects is applied for the specification. Table 3.5 shows there is little evidence of displacement effects of native RNs in response to the influx of foreign-educated RNs, except for the youngest native nurses. The magnitude of the coefficient for U.S.-educated nurses aged 25 to 29 suggests that for every 10 foreign-educated nurses that enter a commuting zone, approximately eight to nine native-educated nurses are employed in the commuting zone. It is intuitive given that foreign-educated nurses are likely to be composed of relatively older nurses, and newly educated native nurses are positively affected by the foreign educated. With respect to education levels, we find positive and significant effects for U.S.-educated RNs with a graduate degree. This result is not

Table 3.5: Displacement effects of foreign-educated RNs on U.S.- educated RNs by age

Dep var: U.S.-educated nurses per 1,000 population									
I. 1990-2015									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-70
Foreign-educated nurses per 1,000 population	0.818**	-0.155	0.387	-0.057	0.394	0.218	0.164	0.331	0.081
	(0.380)	(0.314)	(0.383)	(0.269)	(0.292)	(0.294)	(0.302)	(0.303)	(0.105)
Observations	700	700	700	700	700	700	700	700	700
R-squared	0.631	0.757	0.727	0.792	0.793	0.847	0.877	0.802	0.709
II. 2000-2015									
	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-70
Foreign-educated nurses per 1,000 population	0.906**	-0.068	0.431	-0.514	-0.047	0.040	0.135	0.569	0.125
	(0.761)	(0.391)	(0.594)	(0.489)	(0.450)	(0.397)	(0.429)	(0.545)	(0.181)
Observations	560	560	560	560	560	560	560	560	560
R-squared	0.497	0.736	0.703	0.783	0.808	0.824	0.863	0.743	0.691

Source: The data is from the U.S. Censuses and the American Community Survey multi-year aggregates.

Note: The dependent variable is the number of full-time employed U.S.- educated nurses aged 25-70 years per 1,000 population by age groups in a commuting zone, and the main independent variable is the number of full-time employed internationally educated nurses aged 25-70 years per 1,000 population in a commuting zone. For the specification, the 2SLS model including controls and fixed effects is applied. The instrument is constructed by using the historical distribution of skilled immigrants, excluding those who are registered nurses, across commuting zones in 1980 to allocate the national inflow of foreign-educated nurses to each commuting zone. All specifications are weighted by the whole population in the commuting zone. Standard errors in parentheses are clustered at the commuting zone level. The symbol * means significant at 10% level, ** at 5% level and *** at 1% level.

Table 3.6: Displacement effects of foreign-educated RNs on U.S.- educated RNs by education

	Dep var: U.S.-educated nurses per 1,000 population									
	I. 1990-2015					II. 2000-2015				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Less than HS	High school	Associate	Bachelor	Graduate	Less than HS	High school	Associate	Bachelor	Graduate
Foreign-educated nurses per 1,000 population	0.019	0.207	0.840	0.158	0.658*	0.024	0.175	0.715	0.041	0.924**
	(0.039)	(0.249)	(0.829)	(0.668)	(0.431)	(0.053)	(0.288)	(0.105)	(0.730)	(0.647)
Observations	700	700	700	700	700	560	560	560	560	560
R-squared	0.421	0.890	0.853	0.900	0.755	0.383	0.878	0.902	0.901	0.724

Source: The data is from the U.S. Censuses and the American Community Survey multi-year aggregates.

Note: The dependent variable is the number of full-time employed U.S.- educated nurses aged 25-70 years per 1,000 population by education levels in a commuting zone, and the main independent variable is the number of full-time employed internationally educated nurses aged 25-70 years per 1,000 population in a commuting zone. For the specification, the 2SLS model including controls and fixed effects is applied. The instrument is constructed by using the historical distribution of skilled immigrants, excluding those who are registered nurses, across commuting zones in 1980 to allocate the national inflow of foreign-educated nurses to each commuting zone. All specifications are weighted by the whole population in the commuting zone. Standard errors in parentheses are clustered at the commuting zone level. The symbol * means significant at 10% level, ** at 5% level and *** at 1% level.

surprising given that native nurses with a graduate degree are specialists who are not direct competitors with foreign-educated nurses who typically have a baccalaureate degree. Differences in educational attainment reduce the competition between U.S.- and foreign-educated workers and add to overall complementarity (Borjas, 2003; Ottaviano and Peri, 2006; Wolla, 2014).

The IV estimates are suggestive of the positive association between foreign-educated nurses and native nurses with area fixed effects, and we investigate the potential role of wages to explain the reason for the effects. It is theoretically expected that the influx of immigrant workers would lower the hourly wages of domestic employees, but empirical findings from the literature on the wage effects of foreign workers on their counterparts show a wide range of inconsistency, from negative to positive effects (Borjas and Katz, 2007; Card and Peri, 2016). The effects of immigration in nursing may differ from those observed for workers in general or in other occupations. In the U.S. nursing market, foreign-educated nurses might earn more than native nurses because they are more willing to work under undesirable conditions such as a night shift. In addition, more than 20 percent of domestic nurses work part-time (see Table 3.1).

Table 3.7 reports two separate results for wages. Panel A investigates the hourly wages of employed native nurses, whereas Panel B focuses on the hourly wages of the restricted sample that includes only full-time employed U.S.-educated nurses. The instrumental variable method is used in both Panel A and Panel B. The estimates are mixed and do not point to a systematic association. Large changes in the magnitudes are not observed between Panel A and Panel B, supporting the view that employment status is not the only driving factor in determining hourly wages. We generally fail to find any significant effect of foreign-educated nurses on native wages, similar to Cortes and Pan (2014) and Kaestner and Kaushal (2012).

Table 3.7: Wage effects of foreign-educated RNs on U.S.-educated RNs

	Dep var: Log (Average hourly wages of U.S.-educated Nurses)									
	I. 1990-2015					II. 2000-2015				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A. Employed US-educated Nurses (2SLS – single instrument)										
Foreign-educated nurses per 1,000 population	0.218***	0.232	-0.141	-0.007	-0.091	0.230***	0.389	-0.119	0.048	-0.091
	(0.046)	(0.169)	(0.135)	(0.084)	(0.242)	(0.049)	(0.290)	(0.173)	(0.119)	(0.242)
Observations	700	700	700	420	280	560	560	560	280	280
R-squared	0.532	0.632	0.931	0.963	0.977	0.361	0.212	0.935	0.971	0.977
B. Full-time Employed U.S.-educated Nurses (2SLS – single instrument)										
Foreign-educated nurses per 1,000 population	0.210***	0.202	-0.096	-0.035	0.025	0.220***	0.342	-0.071	-0.010	0.025
	(0.039)	(0.142)	(0.109)	(0.081)	(0.161)	(0.041)	(0.249)	(0.134)	(0.129)	(0.161)
Observations	700	700	700	420	280	560	560	560	280	280
R-squared	0.587	0.691	0.941	0.960	0.978	0.420	0.332	0.942	0.966	0.978
Controls		X	X	X	X		X	X	X	X
Year FE	X	X	X	X	X	X	X	X	X	X
Area FE			X	X	X			X	X	X
Pre-Great Recession				X					X	
Great Recession and Post-Great Recession					X					X

Source: The data is from the U.S. Censuses and the American Community Survey multi-year aggregates.

Note: The dependent variable is the average log hourly wages of native nurses in a commuting zone, and the main independent variable is the number of full-time employed internationally educated nurses aged 25-70 years per 1,000 population in a commuting zone. Wages are deflated so that they are in terms of the 1999 wages using the BLS CPI-U. Panel A examines the hourly wages of employed USTNs (both part-time and full-time) while Panel B focuses on the hourly wages of only full-time employed USTNs. For the 2SLS regressions, the instrument is constructed by using the historical distribution of skilled immigrants, excluding those who are registered nurses, across commuting zones in 1980 to allocate the national inflow of foreign-educated nurses to each commuting zone. All specifications are weighted by the whole population in the commuting zone. Standard errors in parentheses are clustered at the commuting zone level. The symbol * means significant at 10% level, ** at 5% level and *** at 1% level.

Overall, these findings suggest that the positive employment effects between U.S.- and foreign-educated nurses within a commuting zone are largely driven by complementarity across age groups and education levels or by no harmful wage effect of the foreign-educated entry into the U.S. nursing markets.

3.5. Methodologies and results by a multiple instrumentation procedure

Although spatial variation in the concentration of immigrants to identify their labor market impact has been commonly used, Jaeger, Ruist, and Stuhler (2018) point out the conventional instrument can still result in biases. The problem stems from the interplay of general equilibrium adjustments and a correlation between immigrants' country of origin and their settlement pattern. They argue that a single instrument may conflate the short-run wage impacts of immigrant arrivals with the long-run responses to previous supply shocks. If the short-term effects are dominated by the long-run effects, it is common to find different signs on the estimates.

To solve this issue, they introduce a “multiple instrumentation” procedure. By limiting the inclusion of past immigrant inflows to one lag, the empirical specification is as follows:

$$(3.3) \quad \left(\frac{USEN}{Population} \right)_{it} = \alpha + \beta_1 \left(\frac{FEN}{Population} \right)_{it} + \beta_2 \left(\frac{FEN}{Population} \right)_{it-1} + \gamma X_{it} + \delta_i + \lambda_t + \varepsilon_{it}$$

where i is the commuting zone, and t is the time period ($t = 2000, 2007, 2010, \text{ and } 2015$)³⁵. β_1 captures the impact of hiring foreign-educated nurses on domestic nurses in the short-run and β_2 captures the longer-term movement towards equilibrium in response to past supply shocks. Two instruments for two regressors are as follows:

³⁵ To avoid a mechanical relationship between the current inflows of foreign-educated nurses at time t and the lagged instrument, the base year should be strictly prior to $t-1$ (Jaeger et al., 2018).

$$(3.4) \quad \sum_c \frac{Skilled\ Immigrants_{ci,1980}}{Skilled\ Immigrants_{c,1980}} * \Delta FEN_{ct,-i}$$

and

$$(3.5) \quad \sum_c \frac{Skilled\ Immigrants_{ci,1980}}{Skilled\ Immigrants_{c,1980}} * \Delta FEN_{ct-1,-i}$$

where c refers to the country of birth, i the commuting zone and t time period ($t = 2000, 2007, 2010$, and 2015). By instrumenting for foreign-educated nurses' inflows and lagged foreign-educated nurses' inflows with the double instruments, the endogeneity of current and past inflows to current and past labor demand shocks are addressed.

Table 3.8 presents the results of the reduced form and first stage regressions. In the reduced form, the endogenous variables are expressed as functions of the exogenous variables. The base year is always 1980. A negative and statistically significant reduced form relationship is found without a commuting zone fixed effect. However, the coefficients in the reduced form regressions become statistically insignificant when area fixed effects are included. We report the Sanderson-Windmeijer multivariate F statistics for testing for weak instruments in a model with multiple endogenous variables. This statistic is informative about the maximum total relative bias of the 2SLS estimator (Sanderson and Windmeijer, 2016). The null hypothesis is rejected, and the first stage F statistics are reasonably large with the full sample. The first stage coefficients are statistically significant, and they help conclude that the past settlement instrument and its lag affect positively recent and past foreign-educated nurses' inflows.

If the composition of national inflows does not vary much over time, the two instruments can be highly correlated; the past settlement instrument and its lag are constructed using the same base period and a difference between two instruments comes only from variation over time in the composition of national inflows. The correlation across 140 commuting zones between the

Table 3.8: Reduced form and first stage results, 2000-2015 (multiple instrumentation)

	(1)	(2)	(3)	(4)	(5)
Reduced Form					
Past Settlement	-1.945** (0.791)	-0.523* (0.285)	0.482 (0.353)	0.753 (1.843)	-0.461 (1.163)
Lagged Past Settlement	-4.004*** (0.838)	-1.218** (0.605)	0.233 (0.631)	0.854 (2.227)	1.915 (2.861)
First Stages					
<i>Dependent variable: Foreign-educated Nurse Inflows</i>					
Past Settlement	0.961*** (0.178)	0.282** (0.171)	0.291** (0.194)	0.474 (0.674)	0.631 (0.570)
Lagged Past Settlement	1.705*** (0.466)	0.180 (0.205)	0.293 (0.256)	0.144 (0.761)	0.215 (1.160)
Sanderson-Windmeijer multivariate F-statistics	36.50 [0.000]	10.96 [0.000]	10.15 [0.000]	6.30 [0.059]	1.05 [0.307]
<i>Dependent variable: Lagged Foreign-educated Nurse Inflows</i>					
Past Settlement	0.420* (0.216)	-0.088 (0.169)	-0.222 (0.216)	0.512 (0.558)	0.170 (0.570)
Lagged Past Settlement	2.046*** (0.527)	0.660*** (0.252)	0.492** (0.329)	0.110* (0.601)	-0.609 (1.239)
Sanderson-Windmeijer multivariate F-statistics	36.80 [0.000]	25.45 [0.000]	26.69 [0.000]	7.37 [0.008]	5.60 [0.458]
Controls		X	X	X	X
Year FE	X	X	X	X	X
Area FE			X	X	X
Pre-Great Recession				X	
Great Recession and Post-Great Recession					X

Source: The data is from the U.S. Censuses and the American Community Survey multi-year aggregates.

Note: The reduced form reports the slope coefficients from an OLS regression of the number of full-time employed native nurses aged 25-70 years per 1,000 population in a commuting zone on both instruments. The dependent variable in the first-stage regressions is the number of full-time employed foreign-educated nurses aged 25-70 years per 1,000 population in a commuting zone. All specifications are weighted by the whole population in the commuting zone. Standard errors in parentheses are clustered at the commuting zone level. p-values of the F-statistics are in brackets. The symbol * means significant at 10% level, ** at 5% level and *** at 1% level.

instrument and its lag in 1980 is 0.661, which Jaeger, Ruist and Stuhler (2018) might classify as moderate rather than high levels of correlation, with the p-value equal to 0.000³⁶. The spatial distribution of immigrant inflows is stable over time, and we estimate separately the short- and long-run effect of foreign-educated nurses on the labor outcomes of native nurses over 2000-2015. In Table 3.9, the second stage results are reported. For comparison, we also present the conventional (single) IV estimates of the effects of foreign-educated nurses' inflows on native nurses in Panel A. The coefficients for the 2000-2015 time periods (Table 3.4) are only displayed to be consistent with the multiple instrument procedure. We then show the estimates of the effects of foreign-educated nurses' inflows and lagged FENs inflows on the supply of U.S.-educated nurses using Equation (3.3) in Panel B.

The model provides predictions on the signs of the coefficients. The coefficients capture the employment impact of newly arrived foreign-educated nurses in the short-run and longer-term displacement effects in response to local shocks. We find that the short-run impact of recent arrivals of foreign-educated nurses on U.S. nurses is positive but insignificant (Column (3) in Panel B). The estimate is less positive than the corresponding conventional IV estimate in Column (3), Panel A, and it is consistent with our expectation that estimates that do not control for the adjustment to past shocks are biased. We also find a negative and insignificant coefficient on the predicted lagged foreign-educated nurses, in keeping with the expectation that this coefficient captures the longer-term adjustment of local nursing market to local supply shocks induced by foreign-educated nurses. The longer-term coefficient is nearly zero and suggests that

³⁶ Jaeger, Ruist, and Stuhler (2018) note the serial correlation from one decade to the next to show the importance of using the 1960 distribution as their base period for the instrument. On page 22, the correlation in country of origin shares between immigrants arriving in the 1960s and those arriving in 1970s is 0.59, whereas the correlations after 1980 are above 0.90.

Table 3.9. Displacement Effects of Foreign-educated RNs on U.S.-educated RNs, 2000-2015

	Dep var: U.S.-educated nurses per 1,000 population				
	(1)	(2)	(3)	(4)	(5)
A. 2SLS (Single Instrument)					
Foreign-educated nurses per 1,000 population	-2.223***	-3.018*	1.878*	0.237	0.970
	(0.476)	(1.734)	(1.355)	(1.495)	(1.487)
R-squared	0.374	0.664	0.931	0.975	0.967
B. 2SLS (Multiple Instruments)					
Foreign-educated nurses per 1,000 population	-1.837*	-2.245	1.389	0.878	0.333
	(1.010)	(1.701)	(0.797)	(1.890)	(1.899)
Lagged foreign-educated nurses per 1,000 population	-1.426*	-1.236*	-0.035	0.068	3.025
	(0.552)	(0.749)	(0.825)	(1.439)	(1.602)
R-squared	0.379	0.658	0.938	0.972	0.938
Controls		X	X	X	X
Year FE	X	X	X	X	X
Area FE			X	X	X
Pre-Great Recession				X	
Great Recession and Post-Great Recession					X

Source: The data is from the U.S. Censuses and the American Community Survey multi-year aggregates.

Note: The dependent variable is the number of full-time employed U.S.-educated nurses aged 25-70 years in the respective age group per 1,000 population in a commuting zone, and the main two independent variables are the number of full-time employed internationally educated nurses aged 25-70 years per 1,000 population in a commuting zone and its lag. For the 2SLS regressions, two instruments are constructed by using the historical distribution of skilled immigrants, excluding those who are registered nurses, across commuting zones in 1980 to allocate the national inflow of foreign-educated nurses to each commuting zone. All specifications are weighted by the whole population in the commuting zone. Standard errors in parentheses are clustered at the commuting zone level. The symbol * means significant at 10% level, ** at 5% level and *** at 1% level.

Table 3.10. Displacement Effects of Foreign-educated RNs on U.S.-educated RNs by Age, 2000-2015

	Dep var: U.S.-educated nurses per 1,000 per population								
	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-70
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A. 2SLS (Single Instrument)									
Foreign-educated nurses per 1,000 population	0.906**	-0.068	0.431	-0.514	-0.047	0.040	0.135	0.569	0.125
	(0.761)	(0.391)	(0.594)	(0.489)	(0.450)	(0.397)	(0.429)	(0.545)	(0.181)
R-squared	0.497	0.736	0.703	0.783	0.808	0.824	0.863	0.743	0.691
B. 2SLS (Multiple Instruments)									
Foreign-educated nurses per 1,000 population	0.533	-0.134	0.673	-0.346	0.003	-0.081	0.323	0.379	0.039
	(0.508)	(0.375)	(0.549)	(0.439)	(0.284)	(0.319)	(0.280)	(0.480)	(0.116)
Lagged foreign-educated nurses per 1,000 population	-0.488*	-0.048	0.175	0.122	0.036	-0.088	0.136	-0.138	-0.062
	(0.308)	(0.219)	(0.295)	(0.145)	(0.211)	(0.206)	(0.221)	(0.181)	(0.088)
R-squared	0.721	0.734	0.648	0.811	0.810	0.824	0.843	0.795	0.721

Source: The data is from the U.S. Censuses and the American Community Survey multi-year aggregates.

Note: The dependent variable is the number of full-time employed U.S.-educated nurses aged 25-70 years in the respective age group per 1,000 population in a commuting zone, and two main independent variables are the number of full-time employed internationally educated nurses aged 25-70 years per 1,000 population in a commuting zone and its lag. For the 2SLS regressions, two instruments are constructed by using the historical distribution of skilled immigrants, excluding those who are registered nurses, across commuting zones in 1980 to allocate the national inflow of foreign-educated nurses to each commuting zone. All specifications are weighted by the whole population in the commuting zone. Standard errors in parentheses are clustered at the commuting zone level. The symbol * means significant at 10% level, ** at 5% level and *** at 1% level.

U.S.-educated nurses are not eventually affected by foreign nurse recruitment in the long-term trend of employment.

We use the double instruments to look at displacement effects for different age and education groups. Results are presented in Panel B of Table 3.10 and Table 3.11, respectively. We find that the impacts of recent influxes of foreign-educated nurses on the youngest native nurses are positive but insignificant, whereas the lagged inflows of foreign-educated nurses that capture the longer-term effects displace U.S.-educated nurses aged 25 to 29. In terms of education levels, inflows of the foreign-educated nurses are positively related with native nurses with a graduate degree in the short run. However, the lagged estimate is negative and insignificant.

We also find little evidence that an inflow of foreign-educated nurses to a commuting zone causes a decline in the observed wages of native nurses. Panel A of Table 3.12 presents the single IV estimates of the effects of foreign-educated nurses' inflows on native wages, and Panel B reports the double IV estimates of the current and the past FENs inflows on the wages of domestic nurses. The coefficients using the preferred specification that controls for time-varying characteristics and fixed effects (Column (3)) are positive and statistically significant in the long run with the full sample.

One caveat that is worth pointing out is that the double estimates are smaller or larger than its corresponding conventional IV estimates in absolute terms. It implies that estimates that do not control for the adjustment to past supply shocks induced by foreign-educated nurses' inflow are biased.

Table 3.11. Displacement effects of the foreign-educated on U.S.-educated RNs by education, 2000-2015

	Dep var: U.S.-educated nurses per 1,000 population		
	Associate	Bachelor	Graduate
	(1)	(2)	(3)
A. 2SLS (Single Instrument)			
Foreign-educated nurses per 1,000 population	0.715 (0.105)	0.041 (0.730)	0.924** (0.647)
R-squared	0.902	0.901	0.724
B. 2SLS (Multiple Instruments)			
Foreign-educated nurses per 1,000 population	0.771 (0.775)	-0.279 (0.939)	0.882* (0.497)
Lagged foreign-educated nurses per 1,000 population	0.041 (0.530)	-0.232 (0.380)	-0.030 (0.310)
R-squared	0.900	0.895	0.734

Source: The data is from the U.S. Censuses and the American Community Survey multi-year aggregates.

Note: The dependent variable is the number of full-time employed U.S.-educated nurses aged 25-70 years in the respective education group per 1,000 population in a commuting zone, and two main independent variables are the number of full-time employed internationally educated nurses aged 25-70 years per 1,000 population in a commuting zone and its lag. For the 2SLS regressions, two instruments are constructed by using the historical distribution of skilled immigrants, excluding those who are registered nurses, across commuting zones in 1980 to allocate the national inflow of foreign-educated nurses to each commuting zone. All specifications are weighted by the whole population in the commuting zone. Standard errors in parentheses are clustered at the commuting zone level. The symbol * means significant at 10% level, ** at 5% level and *** at 1% level.

Table 3.12. Wage effects of foreign-educated RNs on U.S.-educated RNs, 2000-2015

	Dep var: Log (Average hourly wages of U.S.-educated nurses)									
	I. Employed USENs					II. Full-time Employed USENs				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A. 2SLS (Single Instrument)										
Foreign-educated nurses per 1,000 population	0.230***	0.389	-0.119	0.048	-0.091	0.220***	0.342	-0.071	-0.010	0.025
	(0.049)	(0.290)	(0.173)	(0.119)	(0.242)	(0.041)	(0.249)	(0.134)	(0.129)	(0.161)
R-squared	0.361	0.512	0.935	0.971	0.977	0.420	0.532	0.942	0.966	0.978
B. 2SLS (Multiple Instruments)										
Foreign-educated nurses per 1,000 population	0.117**	0.198	0.043	0.008	-0.149	0.107***	0.169	0.105	0.060	-0.034
	(0.046)	(0.303)	(0.100)	(0.154)	(0.372)	(0.038)	(0.261)	(0.082)	(0.157)	(0.304)
Lagged foreign-educated nurses per 1,000 population	0.125**	0.306*	0.118**	-0.041	-0.277	0.125***	0.277*	0.127**	0.071	-0.278
	(0.051)	(0.184)	(0.060)	(0.108)	(0.694)	(0.046)	(0.161)	(0.061)	(0.120)	(0.160)
R-squared	0.371	0.470	0.950	0.971	0.935	0.433	0.411	0.939	0.966	0.932
Controls		X	X	X	X		X	X	X	X
Year FE	X	X	X	X	X	X	X	X	X	X
Area FE			X	X	X			X	X	X
Pre-Great Recession				X					X	
Great Recession and Post-Great Recession					X					X

Source: The data is from the U.S. Censuses and the American Community Survey multi-year aggregates.

Note: The dependent variable is the average log hourly wages of native nurses in a commuting zone, and two main independent variables are the number of full-time employed internationally educated nurses aged 25-70 years per 1,000 population in a commuting zone and its lag. Wages are deflated so that they are in terms of the 1999 wages using the BLS CPI-U. For the 2SLS regressions, the double instruments are constructed by using the historical distribution of skilled immigrants, excluding those who are registered nurses, across commuting zones in 1980 to allocate the national inflow of foreign-educated nurses to each commuting zone. All specifications are weighted by the whole population in the commuting zone. Standard errors in parentheses are clustered at the commuting zone level. The symbol * means significant at 10% level, ** at 5% level and *** at 1% level.

3.6. Discussion and conclusion

This paper has investigated the employment responses of native nurses to a change in the number of employed foreign-educated nurses in the United States. Using the 1980-2000 U.S. Censuses and the American Community Survey multi-year aggregates obtained from Integrated Public Use Microdata Series (Ruggles et al., 2017), the paper quantifies the displacement effects.

To examine the causal effects of migration of foreign-educated nurses on native nurses, the analysis first closely follows a simple linear model that Cortes and Pan (2014) employed. Because local markets are not closed and immigrants tend to settle in locations where previous waves have settled, the instrumental variable strategy, as well as ordinary least squares are implemented. The instrument for the number of full-time employed foreign-educated nurses is the interaction between the 1980 distribution of skilled immigrants by country across commuting zones and the national inflow of newly arrived foreign-educated nurses from the different countries. Cortes and Pan (2014) use the historical share of skilled immigrants and the stock of foreign-born nurses from the same country of origin in a given metropolitan statistical area to construct their instrument.

We fail to find displacement effect of an inflow of foreign-educated nurses on the number of native nurses by the OLS and IV methodology with a single instrument. On the contrary, the IV estimates rather indicate that an increase in the number of native nurses might stem from an increase in foreign nurses educated abroad. The difference between our results for displacement and those of Cortes and Pan (2014) stem from the differences in definitions of immigrant nurses, with us using foreign educated and Cortes and Pan (2014) using foreign born. We explore which native groups are most affected by foreign-educated nurses and observe that newly educated U.S. nurses and the U.S.-educated nurses with the highest education level benefit from

complementarity. Turning to the possible factors that might drive employment effects of native nurses, we find little evidence that a change in native wages causes the positive effects.

The spatial distribution of immigrant inflows tends to be stable over time, and a single instrument is highly and serially correlated with responses to previous labor shocks. Thus, a multiple instrumentation procedure introduced by Jaeger, Ruist, and Stuhler (2018) is implemented to separate the short- and long-term responses to recent foreign-educated nurses' inflows. Under the preferred model specification with the double instruments, we find no significant effect of foreign-educated nurses on the employment of native nurses in both the short- and long-runs. We find marginally significant displacement effects on the youngest native nurses in the long run, whereas highly educated natives are positively affected by the foreign-educated nurses in the short run. We fail to find the negative impact of foreign-educated nurses on native wages but observe larger in magnitude double IV estimates than the corresponding single IV estimates, which implies that researchers should control for dynamic adjustments to past local shocks.

Taken together, the findings suggest that relying on foreign-educated nurses to fill gaps in the U.S. health care workforce does not harm the employment of native nurses, but rather the entry of foreign-educated nurses to the U.S. nursing market causes an increase in native wages in the long run.

3.7. Chapter 3 references

- Aiken, Linda, Sean Clarke, Robyn Cheung, Douglas Sloane, and Jeffrey Silber. 2003. "Educational levels of hospital nurses and surgical patient mortality." *Jama* 290(12): 1617-1623.
- Aiken, Linda, Sean Clarke, Douglas Sloane, Julie Sochalski, Reinhard Busse, Heather Clarke, Phyllis Giovannetti, Jennifer Hunt, Anne Marie Rafferty, and Judith Shamian. 2017. "Nurses' reports on hospital care in five countries." *Health affairs*.
- Altonji, Joseph G. and David Card. 1991. "The Effects of Immigration on the Labor Market Outcomes of Less-skilled Natives." *Immigration, Trade, and the Labor Market*. University of Chicago Press, 201-234. URL: <http://www.nber.org/chapters/c11773>.
- Arends-Kuenning, Mary. 2006. "The Balance of Care: Trends in the Wages and Employment of Immigrant Nurses in the US between 1990 and 2000." *Globalizations* 3(3): 333-348.
- Anderson, Stuart. 2019. "Latest Data Show H-1B Visas being Denied at High Rates." *Forbes*. Retrieved from <https://www.forbes.com/sites/stuartanderson/2019/10/28/latest-data-show-h-1b-visas-being-denied-at-high-rates/#399897a454c3>
- Autor, David, and David Dorn. 2013. "The growth of low-skill service jobs and the polarization of the US labor market". *American Economic Review* 103(5): 1553-97.
- Benson, Alan. 2012. "Labor market trends among registered nurses: 2008-2011". *Policy, Politics, & Nursing Practice* 13(4): 205-213.
- Borjas, George. 2003. *The Labor Demand Curve is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market*. Working Paper 9755. National Bureau of Economic Research. doi: 10.3386/w9755. URL: <http://www.nber.org/papers/w9755>.
- Borjas, George. 2016. "Yes, Immigration Hurts American Workers". *Politico Magazine*. Retrieved from <https://www.politico.com/magazine/story/2016/09/trump-clinton-immigration-economy-unemployment-jobs-214216>
- Borjas, George, and Lawrence Katz. 2007. "The evolution of the Mexican-born workforce in the United States". *Mexican immigration to the United States*. University of Chicago Press, pp. 13-56.
- Brush, Barbara, Julie Sochalski, and Anne Berger. 2004. "Imported care: Recruiting foreign nurses to US health care facilities". *Health Affairs* 23(3): 78-87.
- Buerhaus, Peter, David Auerbach, and Douglas Staiger. 2009. "The recent surge in nurse employment: Causes and implications". *Health Affairs* 28(4): w657-w668.

- Buerhaus, Peter, Lucy Skinner, David Auerbach, and Douglas Staiger. 2017. "State of the registered nurse workforce as a new era of health reform emerges". *Nursing Economics* 35(5): 229-237.
- Cadena, Brian, and Brian Kovak. 2016. "Immigrants equilibrate local labor markets: Evidence from the Great Recession". *American Economic Journal: Applied Economics* 8(1): 257-90.
- Card, David. 2001. "Immigrant inflows, native outflows, and the local labor market impacts of higher immigration". *Journal of Labor Economics* 19(1): 22-64.
- Card, David, and Ethan Lewis. 2007. "The diffusion of Mexican immigrants during the 1990s: Explanations and impacts". *Mexican immigration to the United States*. University of Chicago Press, pp. 193-228.
- Card, David, and Giovanni Peri. 2016. "Immigration economics by George J. Borjas: a review essay". *Journal of Economic Literature* 54(4): 1333-49.
- Cortes, Patricia, and Jessica Pan. 2014. "Foreign nurse importation and the supply of native Nurses". *Journal of Health Economics* (37): 164-180.
- Cortes, Patricia, and Jessica Pan. 2015. "The relative quality of foreign-educated nurses in the United States". *Journal of Human Resources* (50)4: 1009-1050.
- Dillender, Marcus, Andrew Friedson, Cong Gian, and Kosali Simon. 2019. "Does the healthcare educational market respond to short-run local demand?" *Economics of Education Review* 73: 101032.
- Economic Policy Institute, EPI. 2018. *Methodology for measuring wages and benefits*. URL: <https://www.epi.org/data/methodology>.
- Hobson, Jeremy. 2019. "How the Trump Administration Has Changed Legal Immigration", interviewed by Muzaffar Chishti, *National Public Radio*, available: <https://www.wbur.org/hereandnow/2019/09/23/trump-administration-legal-immigration>
- Jaeger, David, Joakim Ruist, and Jan Stuhler. 2018. *Shift-Share Instruments and the Impact of Immigration*. Working Paper 24285. National Bureau of Economic Research. doi: 10.3386/w24285. URL: <http://www.nber.org/papers/w24285>.
- Kaestner, Robert, and Neeraj Kaushal. 2012. "Effect of immigrant nurses on labor market outcomes of US nurses". *Journal of urban economics* 71(2): 219-229.
- Kline, Donna. 2003. Push and pull factors in international nurse migration". *Journal of nursing scholarship* 35(2): 107-111.

- Glaessel-Brown, Eleanor. 1998. "Use of immigration policy to manage nursing shortages". *Journal of Nursing Scholarship* 30(4): 323-327.
- Lafer, Gordon. 2005. "Hospital speedups and the fiction of a nursing shortage". *Labor Studies Journal* 30(1): 27-46.
- Lovell, Vicky. 2006. *Solving the Nursing Shortage Through Higher Wages: Institute for Women's Policy Research*. Institute for Women's Policy Research.
- Martineau, Tim, Karola Decker, and Peter Bundred. 2002. "Briefing note on international migration of health professionals: levelling the playing field for developing country health systems". *Liverpool: Liverpool School of Tropical Medicine*.
- McDonald, James Ted, Rebecca Ma Valenzuela, et al. 2012. "Why Filipino migrants remit? Evidence from a home-host country matched sample". *Monash University, Department of Economics, Discussion Papers* 9/12.
- Munshi, Kaivan. 2003. "Networks in the modern economy: Mexican migrants in the US labor market". *The Quarterly Journal of Economics* 118(2): 549-599.
- National Foundation for American Policy. 2019. State department visa refusals in FY 2018 for immigrants and nonimmigrants. Retrieved from <https://nfap.com/wp-content/uploads/2019/03/State-Department-Visa-Refusals-In-FY-2018.NFAP-Policy-Brief.March-2019.pdf>
- Ortman, Jennifer, Victoria Velko, Howard Hogan, et al. 2014. *An aging nation: the older population in the United States*. United States Census Bureau, Economics and Statistics Administration, US Department of Commerce.
- Ottaviano, Gianmarco, and Giovanni Peri. 2006. *Rethinking the Effects of Immigration on Wages*. Working Paper 12497. National Bureau of Economic Research. doi: 10.3386/w12497. URL: <http://www.nber.org/papers/w12497>.
- Pierce, Sarah. 2019. "Immigration-related policy changes in the first two years of the Trump administration". *Washington DC, Migration Policy Institute*, available: www.migrationpolicy.org/research/immigration-policy-changes-two-years-trump-administration.
- Ruggles, Steven, Katie Genadek, Ronald Goeken, Josiah Grover, and Matthew Sobek. 2017. *Integrated Public Use Microdata Series: Version 7.0 [dataset]*. URL: <https://doi.org/10.18128/D010.V7.0>.
- Sanderson, Eleanor, and Frank Windmeijer. 2016. "A weak instrument F-test in linear IV models with multiple endogenous variables". *Journal of Econometrics* 190(2): 212-221. URL: <https://EconPapers.repec.org/RePEc:eee:econom:v:190:y:2016:i:2:p:212-221>.

- Schumacher, Edward. 2011. "Foreign-born nurses in the US labor market". *Health economics* 20(3): 362-378.
- Stock, James, and Motohiro Yogo. 2002. Testing for weak instruments in linear IV regression.
- Tsitouras, Diomedes, and Maria Pabon Lopez. 2009. "Flatlining: how the reluctance to embrace immigrant nurses is mortally wounding the US health care system". *J. Health Care L. & Pol'y* 12: 235.
- Trucios-Haynes, Enid. 2002. "Temporary workers and future immigration policy conflicts: Protecting US workers and satisfying the demand for global human capital". *Immigr. & Nat'lity L. Rev.* 23: 357.
- Wolla, Scott. 2014. "The Economics of Immigration: A Story of Substitutes and Complements". *Page One Economics*.
- Yang, Dean, and Claudia Martinez. 2006. "Remittances and poverty in migrants' home areas: Evidence from the Philippines". *International Migration, Remittances and the Brain Drain Washington DC: Worl.*

APPENDIX A: SUPPLEMENTARY MATERIAL FOR CHAPTER 1

Table A.1. Baseline data comparison between the samples of those who decided to provide biomarkers and those who decided against it

	Those who provided biomarkers (N=20,744)	Those who did not provide biomarkers (N=9,094)
Has Diabetes	22.54%	21.19%
<i>Race/Ethnicity</i>		
NHW	64.77%	57.80%
NHB	18.39%	22.42%
Hispanics	13.31%	14.14%
NHO	3.53%	5.64%
Age (mean)	65.33	64.38
Female	57.41%	54.66%
<i>Educational attainment</i>		
Less than HS	21.77%	24.94%
High School (HS)	31.71%	30.16%
Above HS	46.51%	44.90%
Married	65.74%	58.59%
<i>Rurality</i>		
Large central/fringe metro counties	51.83%	55.47%
Medium/small metro counties	27.73%	25.96%
Nonmetro (rural) counties	20.44%	18.57%

Abbreviations: NHW = non-Hispanic Whites; NHB = non-Hispanic Blacks; NHO = non-Hispanic other races

Table A.2. Baseline data comparison between the samples of those who provided their year of diagnosis and those who did not report their year of diagnosis among people with previously diagnosed diabetes

	Those who provided their year of diagnosis (N=1,277)	Those who did not report their year of diagnosis (N=1,429)
<i>Race/Ethnicity</i>		
NHW	52.39%	62.14%
NHB	25.92%	20.36%
Hispanics	18.40%	14.70%
NHO	3.29%	3.50%
Age (mean)	64.16	72.63
Female	54.42%	57.03%
<i>Education attainment</i>		
Less than HS	27.09%	28.62%
High School (HS)	33.20%	34.92%
Above HS	39.70%	38.00%
Married	66.41%	57.87%
<i>Rurality</i>		
Large central/fringe counties	52.70%	44.16%
Medium/small counties	29.76%	32.12%
Nonmetro (rural) counties	17.54%	23.72%

Abbreviations: NHW = non-Hispanic Whites; NHB = non-Hispanic Blacks; NHO = non-Hispanic other races

Effective diabetes management among individuals with undiagnosed diabetes

Several recent studies examine racial and ethnic differences in awareness of chronic disease, with some results related to hypertension showing greater awareness among minorities compared to non-Hispanic White Americans. For example, Ong et al. (2007) found that awareness of hypertension is higher among African Americans than non-Hispanic Whites. If these results are consistently applied to diabetes, minorities are less likely to manage their diabetes than non-Hispanic White Americans because they do not know their diabetes status.

To assess if there are racial and ethnic differences in adequate diabetes management among individuals with undiagnosed diabetes, I analyzed how well those people maintain their A1c values below the threshold of 6.5% in the follow-up. I estimated the logistic equation used for effective diabetes management among people with previously known diagnosed diabetes. In these analyses, I only included 381 people whose blood sugar is greater than or equal to 6.5% but do not self-reported doctor's diagnosis in the baseline (2006/2008). When I included individuals with newly undiagnosed diabetes, the results were consistent with the analyses without people with newly undiagnosed diabetes.

$$Y_{ijt_f} = \beta_0 + \beta_1 \text{Indiv}_{ijt_b} + \beta_2 \text{County}_{jt_b} + W_{it_f} + \varepsilon_{ijt}$$

The dependent variable is equal to 1 when an individual with undiagnosed diabetes i living in j county in the baseline has a A1c value below the threshold of 6.5% in the follow-up. If an individual with undiagnosed diabetes in the baseline exceeds his/her A1c values of 6.5% in the follow-up, the dependent variable is equal to 0, indicating inadequate diabetes management.

Consistent with previous studies that minority groups would be less likely to manage their A1c level in the follow-up compared to non-Hispanic White Americans are, I found statistically significant differences between non-Hispanic Blacks with undiagnosed diabetes and

non-Hispanic Whites with undiagnosed diabetes in effective diabetes management. Controlling for only individual-level demographic factors and residency of the county in the baseline shows that the average Hispanic American with undiagnosed diabetes had about 46% lower odds of having effective diabetes management compared to the average non-Hispanic Whites with undiagnosed diabetes. However, statistical significance is eliminated with the introduction of other individual-level characteristics.

The Health and Retirement Survey is committed to notify the respondents about biomarker results. The notification letter mailed within about a month reports all biomarker levels, the recommended normal ranges, and suggestions to see a physician or other health professional to recheck the results and discuss them with a doctor when the biomarkers were outside normal range. Thus, some people could know that they have diabetes from participating in biomarker survey, and it may explain why I found no statistically significant difference between minorities and non-Hispanic Whites. Some people have known they have diabetes because they have taken blood test and notified their results, but do not pay attention to lower their A1c values in the follow-up. Another empirical challenge in estimating racial and ethnic disparities in awareness is individual-level factors such as disease severity may confound an observed association between race and ethnicity and awareness. For many conditions, minorities are only diagnosed or aware of their status when the disease becomes severe. Unfortunately, individuals with undiagnosed diabetes do not take any medical treatments, and I could not determine the degree of severity for respondents with undiagnosed diabetes.

For individuals with undiagnosed diabetes, one of the most important predictors to cause effective diabetes management is county-level Health Professional Shortage Area (HPSA). The average resident with undiagnosed diabetes in a Health Professional Shortage Area was 37.2%

less likely to manage his/her blood sugar than the average individual with undiagnosed diabetes living in a non-HPSA was.

Table A.3. Odds Ratio of Reporting Effective Diabetes Management among Individuals with Undiagnosed Diabetes

	(1)	(2)	(3)	(4)
<i>Rurality (ref=Large central/fringe counties)</i>				
Medium/small counties (Baseline)	0.951 (0.201)	0.926 (0.212)	0.889 (0.207)	0.893 (0.210)
Rural counties (Baseline)	1.116 (0.298)	1.143 (0.309)	1.154 (0.317)	1.317 (0.386)
<i>Race/Ethnicity</i>				
Hispanics	0.540** (0.152)	0.615 (0.198)	0.616 (0.202)	0.670 (0.221)
NHB	0.606** (0.141)	0.588** (0.143)	0.574** (0.141)	0.626* (0.156)
NHO	0.507 (0.244)	0.542 (0.271)	0.508 (0.258)	0.487 (0.250)
Female (ref=male)	1.674*** (0.315)	1.829*** (0.365)	1.733*** (0.351)	1.744*** (0.355)
Individual-level SES and health behavior (Baseline)				
<i>Education (ref=HS)</i>				
Less than HS		1.129 (0.303)	1.115 (0.302)	1.117 (0.304)
Above HS		1.258 (0.287)	1.260 (0.291)	1.260 (0.293)
Married		1.216 (0.252)	1.252 (0.263)	1.227 (0.261)
Has Medicare		1.143 (0.425)	1.143 (0.429)	1.148 (0.431)
Visited a doctor		1.886** (0.556)	1.535 (0.469)	1.535 (0.472)
<i>Smoking status (ref=Never smoked)</i>				
Former smoker		1.437* (0.294)	1.399 (0.289)	1.449 (0.302)
Current smoker		1.793 (0.569)	1.935 (0.632)	2.215 (0.745)
Overweight		1.265 (0.352)	1.224 (0.347)	1.240 (0.355)

Table A.3. Odds Ratio of Reporting Effective Diabetes Management among Individuals with Undiagnosed Diabetes (cont)

Being obese	0.981 (0.258)	0.946 (0.254)	1.000 (0.271)
Severity of diabetes (Baseline)			
A1c level		0.732*** (0.068)	0.725*** (0.068)
County-level contextual factors (Baseline)			
Quality of Health Care			1.413 (0.362)
HPSA			0.628** (0.139)
Wave FE	X	X	X

Abbreviations: NHW = non-Hispanic Whites; NHB = non-Hispanic Blacks; NHO = non-Hispanic other races; HPSA = Health Professional Shortage Area

Note: All models include dummies for age groups (51-59, 60-64, 65-69, 70-74, and 75 and over), and the four regions in the U.S. Census, with South as the reference category.

Note: Standard errors in parentheses are clustered at the individual level. Significance is indicated by the symbol * at the 10% level, ** at the 5% level, and *** at the 1% level.

APPENDIX B: SUPPLEMENTARY MATERIAL FOR CHAPTER 3

Table B.1. Lists of 140 commuting zones

CZ	State	Metro Status	CZ	State	Metro Status
302	Tennessee	Mixed	19600	New Jersey	Metro
401	North Carolina	Mixed	19700	Pennsylvania	Metro
500	North Carolina	Mixed	19800	Delaware	Metro
900	North Carolina	Mixed	19902	Maryland	Mixed
1100	North Carolina	Mixed	20001	Maine	Mixed
1203	North Carolina	Mixed	20003	Maine	Non-metro
1701	North Carolina	Mixed	20100	Maine	Mixed
1900	North Carolina	Mixed	20200	Vermont	Mixed
2000	Virginia	Mixed	20302	Vermont	Non-metro
2400	Virginia	Metro	20401	Rhode Island	Metro
2500	Virginia	Mixed	20500	Massachusetts	Metro
2700	Mississippi	Mixed	20600	New Hampshire	Mixed
3003	Mississippi	Mixed	20800	Massachusetts	Metro
3300	Louisiana	Mixed	20901	Connecticut	Mixed
3901	Louisiana	Mixed	21501	Minnesota	Metro
4200	Arkansas	Mixed	23100	Wisconsin	Mixed
5202	Tennessee	Metro	23500	Illinois	Mixed
5600	Tennessee	Metro	23801	Illinois	Metro
6000	Alabama	Mixed	23900	Illinois	Mixed
6700	Florida	Metro	24100	Wisconsin	Mixed
6800	Florida	Mixed	24300	Illinois	Metro
6900	Florida	Mixed	24400	Illinois	Mixed
7000	Florida	Mixed	24701	Missouri	Mixed
7100	Florida	Mixed	24802	Illinois	Mixed
7200	Florida	Metro	28401	Colorado	Mixed
7300	Florida	Metro	28900	Colorado	Metro
7400	Florida	Mixed	29301	Kansas	Mixed
7500	Florida	Mixed	29502	Missouri	Mixed
7600	Florida	Mixed	30601	Texas	Mixed
7800	Florida	Mixed	30802	Texas	Mixed
7900	Florida	Mixed	31201	Texas	Mixed
8202	South Carolina	Mixed	31301	Texas	Mixed
8401	Georgia	Mixed	31401	Texas	Mixed
8800	Georgia	Mixed	31600	Texas	Mixed
8900	Georgia	Mixed	31700	Texas	Mixed
9100	Georgia	Mixed	31800	Texas	Mixed
9400	Georgia	Mixed	31900	Texas	Mixed

Table B.1. Lists of 140 commuting zones (cont)

9701	Georgia	Mixed	32000	Texas	Mixed
9900	Florida	Mixed	32100	Texas	Mixed
10900	Florida	Mixed	32201	Texas	Non-metro
11101	Alabama	Mixed	32801	Texas	Mixed
11302	Maryland	Metro	32900	Texas	Metro
11304	Maryland	Mixed	33000	Texas	Mixed
11600	Michigan	Mixed	33100	Texas	Mixed
11700	Michigan	Metro	33300	Texas	Mixed
11900	Michigan	Mixed	33803	Oklahoma	Mixed
12200	Michigan	Mixed	34102	Alaska	Mixed
12501	Ohio	Mixed	34701	Hawaii	Metro
12701	Ohio	Mixed	34901	New Mexico	Mixed
13501	Ohio	Mixed	35001	Arizona	Mixed
13600	Indiana	Mixed	35100	Arizona	Mixed
14200	Indiana	Mixed	35600	Hawaii	Non-metro
14900	Indiana	Metro	36100	Utah	Mixed
15200	Ohio	Mixed	37000	California	Mixed
15900	Ohio	Mixed	37100	California	Metro
16300	Pennsylvania	Mixed	37200	California	Metro
16400	Ohio	Mixed	37300	California	Mixed
16600	Virginia	Mixed	37400	California	Mixed
17400	Pennsylvania	Mixed	37500	California	Metro
17600	Virginia	Mixed	37604	Nevada	Mixed
17700	New York	Mixed	37700	California	Mixed
17800	New York	Non-metro	37800	California	Metro
17900	New York	Mixed	37901	Nevada	Mixed
18000	New York	Mixed	38000	California	Metro
18600	New York	Mixed	38200	California	Metro
18800	Pennsylvania	Mixed	38300	California	Mixed
19100	Pennsylvania	Mixed	38601	Washington	Mixed
19300	New York	Mixed	38801	Oregon	Metro
19400	New York	Metro	38901	Oregon	Mixed
19500	New Jersey	Metro	39400	Washington	Mixed

Note: Metro status indicates whether the individual resided within a metropolitan area and, for individuals in metropolitan areas, whether the individual resided within or outside of a central/principal city. If a commuting zone lies only partially within a metropolitan area or central/principal city, then its metro status is mixed.

Table B.2. Numbers of commuting zones with few nurses

		1990	2000	2007	2010	2015
USEN	Zero	0	0	0	0	0
	>0 & <5	5	5	0	3	22
	>0 & <10	5	9	3	3	39
FEN	Zero	215	180	197	193	217
	>0 & <5	40	41	32	36	85
	>0 & <10	66	63	50	48	109

Note: The data is from the U.S. Censuses and the American Community Survey multi-year aggregates. The samples are restricted to employed registered nurses aged 25-70 years old. The table show the numbers of commuting zones with few nurses across the United States.