The Effect of Rurality on Mental and Physical Health

Steven Stern, Elizabeth Merwin, Emily Hauenstein, Ivora Hinton, Virginia Rovnyak, Melvin Wilson, Ishan Williams, and Irma Mahone*

April 8, 2008

1 Introduction

Any time researchers choose to conduct a study on any aspect of rural health care, they are faced with a difficult methodological choice regarding the operational definition of *rural* to use in their study. While this seems to many who are not familiar with rural research to be a straightforward question (and often naive rural researchers expect to find a single answer that is commonly agreed upon by experts in rural research), to date there is no single answer that is based on a consensus and supported by scientific evidence. Coburn, et al. (2007) states that "there is no single, universally preferred definition of rural that serves all policy purposes." One goal of this paper is to provide a method useful in informing the choice of rural definitions for a specific research study. This methods paper presents a systematic evaluation of the impact of the choice of rural definition on results. It was stimulated by the need to select a rural definition to use in related papers on the impact of community resources on mental and general health outcomes in other research by the authors. A national dataset, the Community Tracking Survey, 2000-2001, includes individual level observations from household interviews. We merge it with county level data reflecting community resources, and we use econometric methods to analyze this multi-level data, accounting for individuals from the same family being included in the dataset. The effect of using four different definitions of rural available for use in county level analysis is presented. A statistical analysis of the impact of the choice of a rural definition on outcomes and on the estimates and significance of explanatory variables in the model is presented and is used to inform the selection of the definition to use in other research. Differences in results for mental health, physical health, and utilization of health care variables are evaluated. The choice of a rural definition is presented and

*We would like to thank Dori Stern for excellent research assistance and Donna Tolson for help with some data collection.

justified using the methodological analyses presented in this paper. Strengths and weaknesses of using county-level community characteristics as compared to data from larger geographic areas, such as Health Services Areas, or from smaller geographic areas, such as census tracts or zip codes, are discussed. Finally, the need for a methods study to guide the use of multi-level geographic data to reflect community characteristics within health care studies is proposed.

2 Literature Review

Examining the variation among rural definitions is important to facilitate understanding the allocation of resources aimed at improving mental and physical health care. Characteristics of the rural context are important because they may represent both adverse and beneficial exposures encountered by the individual that affect health outcomes (Diez-Roux, et al. 2001; Bond Huie, Hummer, and Rogers 2002). Selecting the appropriate measure of rurality, however is not a trivial matter (Jordan and Hargrove 1987; de la Torre, Fickenscher, and Luft 1991; Ricketts and Johnson-Webb 1997; Ricketts, Johnson-Webb, and Taylor 1998; Hart, Larsen, and Lishner 2005). Each of several definitions of rural places introduce biases about what constitute key characteristics of rurality and pose challenges to accurately assessing the unique contributions of rurality to health services utilization and outcomes. Further, these selection criteria profoundly affect the sampling frame, including and excluding respondents dependent upon the rural definition used (Hart et al., 2005). In this methodological paper, we examine the effect of three commonly used measures of rurality - population size, adjacency to an urban area, and a continuous measure of rurality on several health and mental health outcomes and on health care utilization.

Definitions of rural include criteria for differentiating characteristics of rural places from urban places, different units of analysis, and methods for establishing "levels" of rurality (de la Torre, Fickenscher, and Luft 1991; Hart, Larsen, and Lishner 2005). Criteria commonly used include size, adjacency to urban places, economic ties to large urban centers, land-based economies, patterns of transportation, and socio-cultural characteristics. Theoretical considerations may direct the researcher to select one set of criteria over another. The unit of analysis often is at the level of the census tract and county, but rural researchers may use zip codes (de la Torre, Fickenscher, and Luft 1991; Fortney, Owen, and Clothier, 1999; Lin, Crawford, and Salmon 2005) or other units of analysis (Lovett et al. 2002; Yiannakoulias et al. 2003; Gessler, et al. 2004; Niggebrugge et al. 2005). Often the unit of analysis selected is determined by the level of specificity of the dataset being used by the investigator and privacy issues which preclude the use of finer units of analysis. With regard to methods, rural definitions may be dichotomous such as the Office of Management and Budget (OMB) Metropolitan Statistical Area (MSA) vs. non-Metropolitan Statistical Area (non-MSA) designations or graduated such as the Department of Agriculture's Rural-Urban Continuum (Butler and Beale, 1994) or Urban Influence Codes (Ghelfi and Parker, 1995). These latter scales divide MSAs and non-MSAs along a continuum of most urban to most rural. Recently, the Rural-Urban Commuting Codes were introduced by the OMB.[1] These codes use census tract or zip codes as the unit of analysis and classify these units based on the commuting patterns between places (Hart, Larsen, and Lishner 2005). It is important to note that, following each census, definitions are adjusted to reflect changes in the rural population and other shifts in characteristics of rural and urban places occurring in the decade previous to the Census.

Various approaches to defining rural exist based on population density, population size, and social characteristics, or remoteness from urban areas with larger health care facilities. Others advocate using a definition that is most suitable for the issue being examined (Hoggart, 1990). Coburn et al. (2007) presents information on many of the different rural definitions and provides references to sources to assist the field in making selections among different definitions of rural. However, even with many rural definitions available to choose from, not only researchers but government agencies continue to struggle with the choice of a rural definition in the dissemination of their research. A limited literature search of recent rural health services research reveals several different ways of defining rural were identified. Also, sometimes papers focus on research conducted in a "rural" population but rural is not specifically defined. This seems particularly true when it appears that other programs, boards, or initiatives are already in existence and were likely created due to the rural nature of the setting or population served or when the main focus of the article was on a specific topic and the setting just happened to be rural. For example, Wong et al. (2006) show that the utilization of Medicaid health services by American Indians/Alaskan Natives relying on the fact that 95% of individuals using the Indian Health Service used "tribal (rural)" programs. It appeared that it was well accepted that the programs served a rural population and *rural* was mainly a descriptor of the population served by a specific program. In other studies, efforts are made to specifically classify areas according to rurality, but sometimes the specific classification criteris is not clear and may represent a well known classification system that has been used or modified. For example, Hynes et al. (2007) conducted a national study that merged Veterans Affairs and Medicare service files for dually eligible individuals and used a zip code level identification of rural referenced to a VA planning document that was based on census 2000 data. A quick search for the document using Google located the planning group but not the document. Although it is probably available, it is not clear from the article which definition of rural within census data that was used.[2] Other researchers use a specific rural definition such as in Hauenstein et al. (2006) who use the Department of Agriculture's Rural-Urban Continuum Codes REF. These examples describe some of the different ways *rural* is being used or defined in recent research.

Government agencies continue to modify existing definitions of *rural* to de-

[1] See http://www.census.gov/geo/www/ua/ua_2k.html for a discussion of how the Census defines urban.

[2] In both of these examples, it would be possible to identify the geographic areas included in the study and develop more information about the rural population of focus.

fine it for the implementation of their programs. The National Center for Health Statistics has used information from three sources to develop the geographic measure they use to facilitate examining urban and rural differences in health care dividing large metropolitan counties into central and fringe sections. The Health Resources and Services Administration (HRSA, 2007) recognized that some counties classified as metropolitan may contain large areas that "may be rural in nature" and present a database that indicates which metropolitan areas contain segments of rural areas which are eligible for rural designation and therefore eligible to apply for HRSA programs limited to rural areas. Because a particular area may be designated as rural under some definitions but not under others, the choice of rural definition is important in research studies. Researchers must choose a definition and apply it to all areas in their study and therefore are interested in how well a definition works for all of the geographic areas in their study. In contrast, many community leaders and individuals are concerned only with the classification of one particular area. The Rural Assistance Center (2007) allows one to enter information on a specific geographic area such as a zip code, mailing address, or county, and information regarding the classification as rural then is presented according to different definitions for the specific area.

3 Methodology

3.1 Dependent and Explanatory Variables

All of the person-specific data come from the 2000-2001Community Tracking Household Survey (CTS).[3] The CTS data are described in more detail in Section 4. We include a set of 26 dependent variables. We can decompose them into 3 groups, described in Tables 1A, 1B, and 1C. The first group consists of self reported mental health variables, the second group consists of self reported physical health variables,[4] and third group consists of self reported measures of health care provision. The tables label each of the dependent variables, specify the method used in estimation (discussed in more detail in Section 3.3), and define the variable.

[3] We use the restricted version of the CTS that includes geographic identifiers to allow merging with county level data.

[4] We excluded some physical health variables available in CTS because they had some very small cell sizes causing estimators not to converge. These include chronic obstructive pulmonary disease, prostate problems, and skin cancer. We also excluded a number of child health variables.

Table 1A: Mental Health Dependent Variables		
Name	Type	Description
DEPRESS	Probit	Whether Told Has Depression
EMOLIM	Probit	Whether Emotions Limit Work
MCS12	OLS	MCS-12 Score[5].
MENVIS	Probit	Whether Had Any Visit to a Mental Health Professional in Past 12 Months

Table 1B: Physical Health Dependent Variables		
Name	Type	Description
ARTHR	Probit	Whether Told Has Arthritis
ASTHMA	Probit	Whether Told Has Asthma
CANCER	Probit	Whether Told Has Cancer
DIABET	Probit	Whether Told Has Diabetes
HEALIM	Probit	Whether Health Limits Moderated Activities
HLTPRB	Probit	Whether Has Visited a Doctor for Health Problems in Past 12 Months
HEART	Probit	Whether Told Has Heart Disease
HYPTEN	Probit	Whether Told Has Hypertension
LIMIT	Probit	Whether Limited at Work or Other Daily Activities by Health in Last 4 Weeks
PCS12	OLS	PCS-12 Score[6].

Table 1C: Health Care Provision Dependent Variables		
Name	Type	Description
DRVIST	Count	# Doctor Visits
ERVIST	Count	# Emergency Room Visits Excluding those Leading to Inpatient Stays
HSPSTA	Count	# Hospital Stays Excluding Birth
HSPEMR	Count	# Hospital Stays Thru Emergency Room
NTHSP	Count	# Nights in Hospital
OUTSRG	Count	# Outpatient Surgeries
PUTOFF	Probit	Whether Put Off Medical Care in Past 12 Months
SURGRI	Count	# Surgeries
SURGOS	Count	# Surgeries with Overnight Stay
UNMET	Probit	Whether Has Unmet Medical Need
Note: The period of time for all count variables in this table is 12 months		

The explanatory variables come from the CTS, the Area Resource File, Cen-

[5] MCS12 is the SF-12 Mental Component Summary, a continuous measure of mental health increasing in better health. See Ware, et al.(1994) for more information.

[6] PCS12 is the SF-12 Physical Component Summary, a continuous measure of physical health increasing in better health. See Ware et al. (1994) for more information.

sus data, and the Census EEO File.[7] Some estimates of the number of providers in each county are based on a special tabulation obtained from the Census Bureau. The explanatory variables fall into four groups: personal characteristics, site dummies, community characteristics, and interactions between personal and county characteristics. The included personal characteristics are described in Table 2A. We include demographic characteristics, financial characteristics, and insurance provision characteristics. The site dummies are 60 dummy variables defining the geographic site of the observation. These control for regional characteristics not captured by the county characteristics.[8] The community characteristics are collected at the county level,[9] and they are described in Table 2B.[10] All variables measuring numbers of people are defined as percentages or as per 10K population rates. Finally, the interactions we include are described in Table 2C. We allow for interactions between personal race[11] and county racial characteristics, personal race and county rural characteristics, and personal race and the county ratio of racial group health professionals to size of racial group.[12]

Table 2A: Explanatory Personal Characteristics	
Demographic Variables	Hispanic, African-American, Age, Female, Education, Speaks Spanish, Smokes
Financial Variables	Employed, Family Income
Insurance Variables	Source of insurance (or none)

[7] Estimates of non-MD provider variables are based on information from a special tabulation of mental health providers and registered nurses in each county by the U.S. Census Bureau through a special tabulation data request and from information in the EEO file.

[8] We estimated all of the models without site dummies as well.

[9] Virginia is divided up into counties and independent cities.

[10] See Section 6 for a discussion of the appropriateness of using county variables as measures of community characteristics.

[11] Throughout the paper, when we use the term "race," we are referring to being African-American or Hispanic.

[12] We experimented with other variables such as hospitals per capita, advanced degree health professionals of many types per capita, proportion of health professionals of a given race or ethnicity, and proportion of health professionals of a given race or ethnicity interacted with personal race or ethnicity.

Table 2B: Explanatory County Variables	
Demographic Variables	Rural, % African-American, % Hispanic, Median Household Income, % Families Living in Poverty, Median Housing Value
Medical Health Professionals Variables	MDs, ObGyns, Child Psychiatrists, Psychiatrists, DOs, Nurse Practitioners, RNs, Psychologists, Counselors, Social Workers, African-American Health Professionals/10K African-Americans, Hispanic Health Professionals/10K Hispanics
Other Medical Resources Variables	Hospital Beds, Community Mental Health Centers, HMO Penetration

Notes:
1) Except when otherwise specified, medical health professional variables and other medical resources variables are measured as rates per 10K Census.
2) See Section 3.2 for the definitions of *rural* used.

Table 2C: Explanatory Interactions Between Personal and County Characteristics		
Personal Characteristic	County Characteristic	
Race	Rural	Hispanic × Rural, African-American × Rural
Race	Race	Hispanic × % Hispanic, African-American × % African-American
Race	Race of Health Professionals	Hispanic × Hispanic Health Professionals/10K Hispanics, African-American × African-American Health Professionals/10K African-Americans

3.2 Proposed Definitions of *Rural*

We use four definitions of *rural*.[13] All but one are based on the Department of Agriculture's nine rural-urban continuum codes (USDA, 1984) (RUCC), defined in Appendix 9.1. The exception is the percentage of the county that is urban; we call this variable throughout *%URBAN*. Coding values that are based on the RUCC are aggregated from nine codes to a smaller number of values corresponding to different notions of *rural*. These are described in Table 3. Our

[13] We put *rural* in italics when we want to denote the concept of ruralness.

BINARY rural variable is the frequently-used metro-nonmetro variable; *ADJAC* focuses on notions of adjacency to urban areas; and *POPSIZE* focuses on notions of urban population size.

Table 3: Aggregated Rural Codes

Variable Name	Aggregation Scheme	Reference Category	Description
BINARY	(1,2,3); (4-9)	(1,2,3)	Binary description of *rural* with urban counties as the reference category
ADJAC	(1,2,3); (4,6,8); (5,7,9)	(1,2,3)	Reference category is metro; (4,6,8) are counties adjacent to metro counties; and (5,7,9) are counties not adjacent to metro counties
POPSIZE	(1,2,3); (4,5); (6,7); (8,9)	(1,2,3)	Reference category is metro; each nonmetro group has less metro population than the group before

3.3 Estimation and Testing Methods

We have three different types of dependent variables: continuous, binary, and count. For the continuous dependent variables, we use ordinary least squares (OLS). For the binary variables, we use probit and allow for correlation of errors across family members.[14] The estimation method is described in detail in Appendix 9.3. For the count variables, we use count methods and allow for correlation of errors across family members.[15] The estimation method is described in detail in Appendix 9.4.

The three *rural* definitions in Table 3, *BINARY*, *ADJAC*, and *POPSIZE*, have 2, 3, and 4 levels, respectively, and they require 1, 2, and 3 indicator variables, respectively. Conditional on a *rural* definition, we are interested in testing the joint significance of the *rural* variable coefficients and testing whether the *rural* variable coefficients are equal. Define β_r as the vector of *rural* coefficients.[16] We want to test $H_0 : \beta_r = 0$ against $H_A : \beta_r \neq 0$; in Section 5, this test will be referred to as the "overall" test. This is a joint test of whether rurality matters (relative to a reference category corresponding to urban

[14] We also estimated the model without correlation across family members.

[15] We also estimate models with person-specific unobserved heterogeneity (Hausman, Hall, and Griliches, 1984) and without any unobserved heterogeneity.

[16] For the purposes of this discussion, we do not include the interactions of *rural* with personal characteristics in β_r.

counties). Also, for those specifications of *rural* that take on $R > 1$ discrete values (not including the reference category), we want to test $H_0 : \beta_{r1} = \beta_{r2} = .. = \beta_{rR}$ against $H_A : \beta_{r1} \neq \beta_{r2} \neq .. \neq \beta_{rR}$. This is a test of whether the data can distinguish between a binary rural/urban county classification and a richer county classification. For example, if the coefficients for the three nonmetro indicators of the *POPSIZE* variable (indicating $RUCC = 4, 5$, $RUCC = 6, 7$, and $RUCC = 8, 9$, respectively) are all equal, then all of the nonmetro counties have the same effect on the dependent variable, and one might just as well use the *BINARY* variable, which groups all nonmetro counties together. In Section 5, this test will be referred to as the "equality" test. This is a test of whether the data can distinguish between a binary rural/urban county classification and a richer county classification. We use Wald tests to perform both sets of tests, and Appendix 9.5 provides details of the test statistics.

4 Data

4.1 Source of Data and Selection Criteria

The two main sources of data for this project are the Community Tracking Survey[17] (CTS) and the Area Resource File[18] (ARF) for 2004. The CTS is a sample of 56343 people from 60 sites across the United States. Each site is a collection of contiguous counties.[19] For each individual, we observe some basic demographic and financial information described in Table 2B, some measures of health outcomes and medical resource use described in Tables 1A, 1B, and 1C, and the county of residence. The CTS collects information on multiple members of households, leading to a need to control for family-specific effects noted in Section 3.3 and discussed in more detail in Appendices 9.3 and 9.4. The ARF is a health planning data set compiled by the Health Resources and Services Administration (HRSA), and it provides information on each county in the United States. Original data are from a large number of other data sources. The original data sources are aggregated to the county level and included in this database. For each county, we observe some demographic and economic information described in Table 2B and measures of supply of health resources also described in Table 2B. The county of residence in the CTS allows us to merge the CTS data with the ARF data. Variables reflecting health professional availability come from the ARF as well as from estimates based on a Special Tabulation of Census data from 2000.

We use a number of selection criteria focused on deleting observations with missing critical variables. We lose 483 observations because of missing education, 2324 because of missing smoking use, 140 because of missing Hispanic status, 257 because of missing race, and 1 because of a missing FIPS code. In

[17] See Center for Studying Health System Change (2003) for more information.
[18] See U.S. Dept. of HHS (2004a,b) for more information.
[19] For example, the Little Rock site consists of Faulkner, Lonoke, Pulaski, and Saline Counties.

total, we lose 3205 observations for missing variables and end up with a sample of 53138 individuals in 30307 discrete families.

4.2 Sample Moments

Given the focus of this paper, we report sample moments only for the rural variables.[20] There are a total of 433 counties in the 60 sites included in the CTS. The distribution of the RUCCs over these counties and in all United States counties are displayed in Table 4. Urban counties are overrepresented in the CTS with 67.4% of CTS counties having a code ≤ 3 while only 34.7% of US counties have a code > 3. Nonrural individuals in the CTS are even more overrepresented with 89.2% having a code ≤ 3.

Table 4: Distribution of RUCCs

Code	Counties in CTS Sites	Individuals in CTS	All Counties in US
1	203	30063	413
2	62	14849	325
3	27	2465	351
4	25	2071	218
5	5	366	105
6	49	1540	609
7	25	1001	450
8	17	372	235
9	20	411	435
Total	433	53138	3141

We are concerned not so much with cell sizes in the CTS but with cell sizes for the aggregated codes described in Table 3 and for the interactions of the codes with other variables described in Table 2C. These are presented in Table 5. One sees that, without interactions, all of the cell sizes are large enough with the smallest being 777 individuals having $RUCC = (8, 9)$ in the *POPSIZE* rural specification. However, once we allow for interactions of *rural* with *Hispanic* and *African-American*, some cell sizes become very small. The worst case is having only 12 observations for Hisp$\times RUCC = (8, 9)$ in the *POPSIZE* rural specification. This essentially means that we would be using only 12 observations to identify the effect of *Hispanic* interacted with $RUCC = (8, 9)$. For example, with a binary variable such as depression, the estimate would rely on the proportion of the 12 individuals who said they were depressed. We think that any cell size less than 100 is problematic and any significantly smaller is not feasible. Thus, the results of Table 5 suggest that there is a binding tradeoff between the richness of the *rural* specification and the inclusion of important interactions.

[20] Other sample moments are available in Appendix 9.2.

Table 5: CTS Cell Sizes					
POPSIZE		*ADJAC*		*BINARY*	
$(4,5)$	2437	$(4,6,8)$	3983	(>3)	5761
$(6,7)$	2541	$(5,7,9)$	1778		
$(8,9)$	783				
Hisp$\times$ $(4,5)$	73	Hisp$\times$ $(4,6,8)$	104	Hisp$\times$ (>3)	163
Hisp$\times$ $(6,7)$	78	Hisp$\times$ $(5,7,9)$	59		
Hisp$\times$ $(8,9)$	12				
AfrAm$\times$ $(4,5)$	175	AfrAm$\times$ $(4,6,8)$	399	AfrAm$\times$ (>3)	609
AfrAm$\times$ $(6,7)$	229	AfrAm$\times$ $(5,7,9)$	210		
AfrAm$\times$ $(8,9)$	205				

Note: Numbers in parentheses are subsets of RUCCs.

The alternative to using aggregated RUCCs is to use the *%URBAN* measure of *rural*. One would expect the *%URBAN* measure to be highly correlated with the RUCCs; counties with high codes (see Appendix 9.1) should be mostly rural. Figure 1 shows that the two variables are, to a great degree, measuring different characteristics of *rural*. The "Mean" curve is the average value of *%URBAN* conditional on each value of the RUCCs, and the "95% conf bound" curves are the mean plus and minus two standard deviations. It is quite obvious that there is not a tight relationship between the two measures. Deviations between the two measures are due to counties like Louisa County in Virginia. Louisa County is 55 miles from Richmond and is considered part of the Richmond MSA. Because of that, it has a RUCC of 1, the same as New York City. On the other hand, it has a population density of 51.5 persons per square mile (while New York City's population density is 26400), and 0% of it is urban (as measured by the *%Urban* code).

5 Results

Each of the models is estimated using the methods described in Section 3.3, and complete results are available from Stern upon request. In this section, we focus on effects of specifying *rural* in different ways. Two of the key issues in choosing a definition of *rural* is whether the estimated effect of *rural* is significant and whether estimated variation in the effects of *rural* are significant. Figure **??** shows the estimated variation in the effect of *rural* for the six different non-MSA RUCCs using *POPSIZE*, *ADJAC*, and *BINARY* from Table 3. The two curves for each *rural* definition define a 95% point-wise confidence interval for each RUCC value excluding the reference category group $(1-3)$. For example, a 95% confidence interval for the coefficient estimate associated with $RUCC = (6,7)$ using the *POPSIZE* specification is $(-0.502, -0.078)$. Note that, for the *POPSIZE* specification, $RUCC = 4$ and $RUCC = 5$ have the same confidence interval, $RUCC = 6$ and $RUCC = 7$ have the same confidence interval, and $RUCC = 8$ and $RUCC = 9$ have the same confidence interval. These

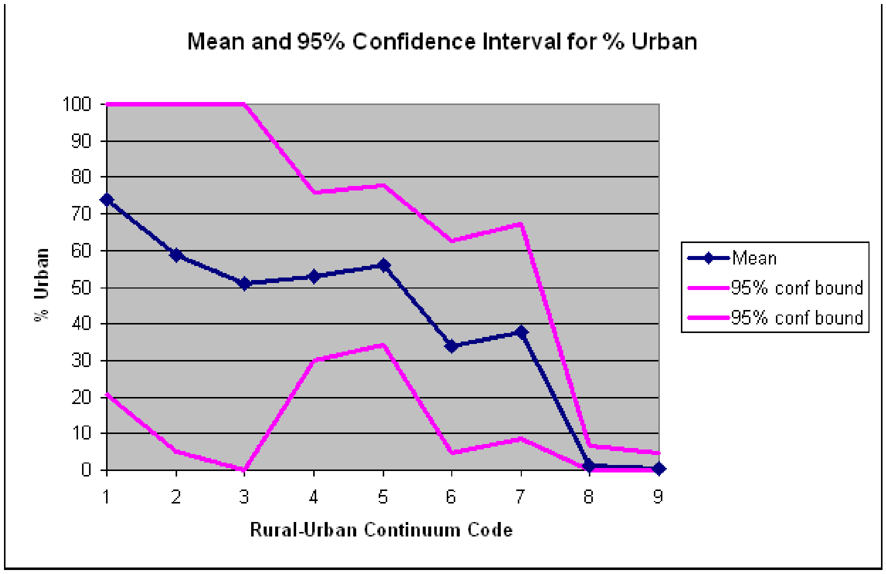

Figure 1: Mean and 95% Confidence Interval for %Urban

equalities are required by the aggregation method implied by the *POPSIZE* specification. An analogous result is true for the *ADJAC* and *BINARY* specifications. Further, note that, although there is variation in the confidence intervals for the three different specifications, all three confidence intervals share a common value of -0.1; i.e., the figure suggests that one would not reject the null hypothesis that the effect of *rural* is the same for all RUCCs greater than 3.

On the other hand, there are some dependent variables where the 95% pointwise confidence intervals look like those in Figure 2. In this case, the data support the (alternative) hypothesis that all three *rural* specifications are actually measuring different features of ruralness. This suggests that the best choice for a *rural* specification varies across dependent variables. However, one must be very careful in choosing a *rural* specification based on analysis implied by Figures **??** and 2. First, it is well known that using such a procedure biases inference. Second, choosing is not straightforward because the specifications are not nested.[21] Finally, the pointwise confidence intervals displayed in Figures **??** and 2 ignore the existence of nonzero covariances among the different *rural* estimates.

[21] See Vuong (1989) for a method to perform non-nested tests.

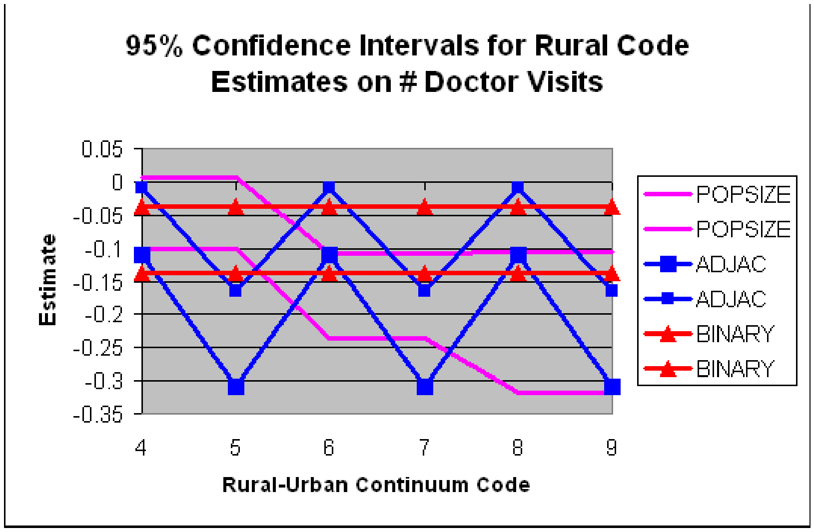

Figure 2: 95% Conf Intvls: Doctor Visits

5.1 Wald Test Results

In fact, the best way to compare estimates from different specifications is with Wald tests described in Section 3.3 and Appendix 9.5. Wald test results are presented in Tables 6A, 6B, and 6C.[22] Table 6A implies that population density is the appropriate measure of *rural* for depression (*DEPRESS*) and mental health visits (*MENVIS*) because two of the measures that change with population density (*BINARY* and *POPSIZE* but *%URBAN* only marginally) are all statistically significant.[23] Furthermore, the χ^2_2 statistic for equality implies rejecting the null hypothesis for *DEPRESS* and is marginally significant for *MENVIS*; i.e. the *POPSIZE* specification is capturing something that *BINARY* cannot capture. However, with only one exception, none of the other χ^2 statistics associated with the mental health dependent variables are significant for either test; the *BINARY* specification is significant only for the overall test for *MCS12*. One should note that, since *DEPRESS* reflects having been told by a medical professional that one has depression, the lower frequency of medical

[22] 5% and 10% critical values for the Wald test statistics are

df	10%	5%
1	2.706	3.841
2	4.605	5.991
3	6.251	7.815

.

[23] Thoughout the paper, when we use the term "statistically significant," we will mean at the 5% level.

professional visits in rural areas[24] may influence the results for *DEPRESS*.

For the physical health dependent variables in Table 6B, the cases where *rural* matters overall are arthritis (*ARTHR*), "health is limiting" (*HEALIM*), and health problems (*HLTPRB*). For *ARTHR* and *HEALIM*, both *BINARY* and *ADJAC* are significant. *ADJAC* is also significant for *HLTPRB* and *HEART*, and *POPSIZE* is significant only for *HEALIM*. *%URBAN* is significant only for *CANCER*. For the equality tests, *POPSIZE* is significant only for *HEALIM*, and *ADJAC* is significant only for *HLTPRB* and *HEART*. For five out of ten of the physical health dependent variables, rural is not significant for any specification.

Table 6A: Wald Test Results for Mental Health Dependent Variables

Name	Overall				Equality	
	BINARY	*ADJAC*	*POPSIZE*	*%URBAN*	*ADJAC*	*POPSIZE*
df	1	2	3	1	1	2
DEPRESS	3.44*	3.78	9.45**	3.50*	0.40	6.23**
EMOLIM	0.23	0.50	1.44	0.16	0.34	1.23
MCS12	4.16**	5.09*	4.78	0.12	1.12	0.38
MENVIS	2.66	4.72*	8.75**	2.76*	2.28	5.90*

Notes:
1) Items with double stars are significant at the 5% level, and items with single stars are significant at the 10% level.
2) See Section 3.3 for a definition of the tests for "Overall" "Equality."

[24] The parameter estimates for the effect of *rural* on *MENVIS* and *DRVIST* are generally negatively related to rurality and statistically significant using either *ADJAC* or *POPSIZE*.

Table 6B: Wald Test Results for Physical Health Dependent Variables						
Name	Overall				Equality	
	BINARY	*ADJAC*	*POPSIZE*	*%URBAN*	*ADJAC*	*POPSIZE*
df	1	2	3	1	1	2
ARTHR	4.74**	7.99**	6.10	1.57	3.36*	0.98
ASTHMA	0.74	0.84	3.33	2.01	0.06	2.47
CANCER	1.46	1.21	1.38	4.03**	0.09	0.06
DIABET	0.40	1.83	4.10	0.24	1.46	3.63
HEALIM	6.31**	6.57**	12.26**	0.63	0.09	6.47**
HLTPRB	2.67	10.75**	5.04	0.15	8.04**	2.32
HEART	0.03	7.14**	4.48	0.57	7.09**	4.48
HYPTEN	1.13	3.13	3.25	0.00	2.05	2.14
LIMIT	0.40	2.40	1.54	0.14	1.96	1.24
PCS12	1.34	3.81	1.98	2.46*	2.43	0.90

Notes:
1) Items with double stars are significant at the 5% level, and items with single stars are significant at the 10% level.
2) See Section 3.3 for a definition of the tests for "Overall" and "Equality."

For the health care provision dependent variables in Table 6C, the cases where rural matters overall are doctor visits (*DRVIST*), emergency room visits leading to a hospital stay (*ERVIST*), hospital stays from the emergency room (*HSPEMR*), outpatient surgeries (*OUTSRG*), "put off medical care" (*PUTOFF*), and surgeries (*SURGRI*). *BINARY* is significant for *DRVIST* and *PUTOFF*; *ADJAC* is significant for *DRVIST*, *ERVIST*, *OUTSRG*, and *PUTOFF*; *POPSIZE* is significant for *DRVIST*, *OUTSRG*, and *SURGRI*; and *%URBAN* is significant for *DRVIST* , *ERVIST*, and whether has unmet medical need (*UNMET*). The dependent variable that have insignificant *rural* effects for all specifications is number of hospital stays excluding birth (*HSPSTA*), number of nights in hospital (*NTHSP*), and surgeries with overnight stays (*SURGOS*). Thus, while there is not strong evidence that *rural* affects either mental or physical health problems, it does have a significant effect on the provision of health care services in most of the cases we tested.[25]

The equality test statistics in Table 6C are also somewhat significant. They are significant for both *ADJAC* and *POPSIZE* for *DRVIST*, *OUTSRG*, and *UNMET*; they are significant for only *ADJAC* for *ERVIST*; and they are significant for only *POPSIZE* for *HSPEMR* and *SURGRI*. These results suggest that the effect of *rural* is subtler than can be captured by a binary rural variable (e.g., *BINARY*), but the important characteristic of *rural* not captured by *BINARY* is different across different dependent variables; for example, for *ERVIST*, *ADJAC* is statistically significant and *POPSIZE* is not, while the

[25] One might consider moving *MENVIS* from Table 6A to Table 6C. This would add one more dependent variable with significant *rural* effects to Table 6C.

opposite is true for *HSPEMR* and *SURGRI*.

Table 6C: Wald Test Results for
Health Care Provision Dependent Variables

Name	Overall				Equality	
	BINARY	*ADJAC*	*POPSIZE*	*%URBAN*	*ADJAC*	*POPSIZE*
df	1	2	3	1	1	2
DRVIST	12.59**	45.85**	34.22**	14.18**	32.00**	21.17**
ERVIST	1.84	17.37**	2.05	6.62**	15.50**	0.27
HSPSTA	2.93*	3.28	3.86	0.96	0.52	0.57
HSPEMR	3.02*	2.99	9.21**	0.18	0.18	6.59**
NTHSP	0.16	0.25	3.17	1.17	0.02	2.80
OUTSRG	2.50	8.27**	22.22**	1.88	5.68**	19.81**
PUTOFF	5.40**	6.20**	7.07*	2.57*	1.03	1.68
SURGRI	1.12	2.48	8.35**	0.91	1.35	7.58**
SURGOS	0.46	1.41	3.19	0.03	1.02	2.65
UNMET	0.16	5.45*	6.56*	7.75**	5.20**	6.48**

Notes:
1) Items with double stars are significant at the 5% level, and items with single stars are significant at the 10% level.
2) See Section 3.3 for a definition of the tests for "Overall" and "Equality."

5.2 Monotonicity Properties of Rural Estimates

One might expect not only tests of overall significance to be informative but also the direction of the effect of increasing ruralness. For example, in Figure 2, using the *ADJAC* specification, we see that, if we ignore interaction terms, then going from an urban county ($RUCC = 1, 2, 3$) to a non-metro county adjacent to an urban area ($RUCC = 4, 6, 8$) decreases the number of doctor visits (*DRVIST*), and going from an $RUCC = 4, 6, 8$ county to a non-metro county not adjacent to an urban area ($RUCC = 5, 7, 9$) further decreases *DRVIST*. Thus, we would say that the effect of *rural* on *DRVIST* is monotone decreasing as we move away geographically from urban areas. On the other hand, using the *POPSIZE* specification for *rural*, we see that the effect of *rural* on *DRVIST* is not monotone in urban population size. In particular, $RUCC = 6, 7$ has a larger negative effect on *DRVIST* than both $RUCC = 4, 5$ or $RUCC = 8, 9$.[26] Whether the nonmonotonicity for the *POPSIZE* specification is problematic depends on whether it should be expected that the effect of *rural* should be monotone. It might be that lack of monotonicity is capturing a real and important effect that could not be captured by a binary indicator of *rural*, or it might be that most of the variation across different counties with respect to rural is noise. The fact

[26] It is meaningless to discuss monotonicity properties for the binary rural variable because all binary variables are monotone. *%URBAN* is mononotone because there are no higher order nonlinear terms for *%URBAN*.

that, for both *ADJAC* and *POPSIZE*, a majority of those cases where the rural estimates are significant overall also reject the equality hypothesis suggests the former is at least partially true.

Tables 7A, 7B, and 7C provide information on the monotonicity properties of *ADJAC* and *POPSIZE* for each of the dependent variables. Table 7A and the *ADJAC* specification suggest that increasing distance from urban areas improves mental health since it decreases *EMOLIM* and *MENVIS* and increases *MCS12*. The *POPSIZE* specification is not monotone for any of the mental health dependent variables. From Table 6A, using *POPSIZE*, only *DEPRESS* and *MENVIS* were significant overall. The estimates for *DEPRESS* using *POPSIZE* are -0.094 for $RUCC = 4,5$, -0.079 for $RUCC = 6,7$, and -0.328 for $RUCC = 8,9$. This suggests that reported depression is significantly less likely in very rural counties ($RUCC = 8,9$) relative to all other counties. The estimates for *MENVIS* were -0.083 for $RUCC = 4,5$, -0.290 for $RUCC = 6,7$, and -0.012 for $RUCC = 8,9$ which does not have any obvious interpretation.

Table 7A: Monotonicity Properties for Mental Health Dependent Variables

Dependent Variable	*ADJAC*	*POPSIZE*
DEPRESS	no monotonicity	no monotonicity
EMOLIM	decreasing	no monotonicity
MCS12	increasing	no monotonicity
MENVIS	decreasing	no monotonicity

Table 7B: Monotonicity Properties for Physical Health Dependent Variables

Dependent Variable	*ADJAC*	*POPSIZE*
ARTHR	decreasing	no monotonicity
ASTHMA	increasing	increasing
CANCER	no monotonicity	no monotonicity
DIABET	no monotonicity	no monotonicity
HEALIM	no monotonicity	no monotonicity
HLTPRB	decreasing	no monotonicity
HEART	no monotonicity	no monotonicity
HYPTEN	decreasing	decreasing
LIMIT	decreasing	no monotonicity
PCS12	increasing	no monotonicity

Table 7B and the *ADJAC* specification suggest that, with the exceptions of *ASTHMA*, physical health either improves with rurality or in not monotone with rurality. In particular, *ARTHR*, *HLTPRB*, *HYPTEN*, and *LIMIT* decrease with distance from urban areas, and *PCS12* increases with distance from urban areas. Table 6B shows that some of those effects are statistically significant. As a group, they paint a picture showing improving health with distance from urban areas. *ARTHR* and *HLTPRB* have significant Wald statistics in

Table 6B and decreasing effects in Table 7B, while *HEART* and *HEALIM* have significant Wald statistics in Table 6B without a monotonicity result. The estimates for *HEART* are 0.018 for $RUCC = 4, 6, 8$ and -0.182 for $RUCC = 5, 7, 9$, and the estimates for *HEALIM* are -0.135 for $RUCC = 4, 6, 8$ and -0.116 for $RUCC = 5, 7, 9$. This again suggests that heart disease is significantly less for non-metro counties not adjacent to metro areas ($RUCC = 5, 7, 9$) relative to all other counties. On the other hand, the *POPSIZE* specification suggests no monotone patterns in physical health other than *HYPTEN* decreasing and *ASTHMA* increasing. Only *HEALIM* had significant *POPSIZE* estimates in Table 6B, and the estimates are -0.178 for $RUCC = 4, 5$, -0.028 for $RUCC = 6, 7$, and -0.111 for $RUCC = 8, 9$.

Table 7C: Monotonicity Properties for Health Care Provision Dependent Variables

Dependent Variable	*ADJAC*	*POPSIZE*
DRVIST	decreasing	decreasing
ERVIST	decreasing	decreasing
HSPSTA	increasing	no monotonicity
HSPEMR	no monotonicity	no monotonicity
NTHSP	increasing	no monotonicity
OUTSRG	decreasing	no monotonicity
PUTOFF	no monotonicity	no monotonicity
SURGRI	decreasing	no monotonicity
SURGOS	increasing	no monotonicity
UNMET	no monotonicity	no monotonicity

Table 7C has a number of monotonicity results, more for *ADJAC* than for *POPSIZE*, but there is no consistent pattern across dependent variables. For *HSPSTA*, *NTHSP* and *SURGOS*, resource use is increasing with ruralness, while for *DRVIST*, *ERVIST*, *OUTSRG*, and *SURGRI*, resource use is decreasing with ruralness. These findings suggest that individuals from rural areas are more likely to use inpatient based care and may overutilize expensive hospital care while using less outpatient care, raising the question of how the lack of access to outpatient care may be resulting in more costly care and care provided in a less desirable setting. This finding suggests that the reduced doctors visits may result in increased hospitalizations. But it does not allow us to determine whether the hospitalization substitutes for care that could be provided through a doctor visit if access were possible. Nor does it allow us to determine if lack of access results in complications and patients being sent later in the course of their illness requiring care only in inpatient settings.

One might look for other systematic patterns in the rural estimates besides monotonicity. While, when using the *POPSIZE* specification, we could find no other patterns for the mental or physical health dependent variables, we did observe a pattern for the utilization variables. In particular, it seemed as if the estimated ratios of the estimates for $RUCC = 6, 7$ over the estimates for $RUCC = 4, 5$ were almost all greater than one and the ratios of the estimates for

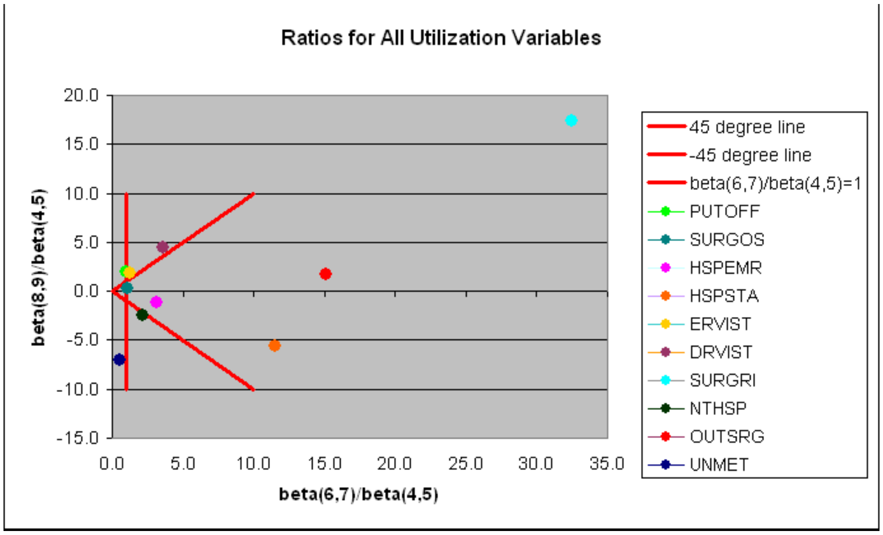

Figure 3: Ratios for All Utilization Variables

RUCC$= 8, 9$ over the estimates for *RUCC*$= 4, 5$ were mostly smaller in absolute value than the ratios for *RUCC*$= 6, 7$ over the estimates for *RUCC*$= 4, 5$. Figure 3 displays the ratios for each of the utilization variables. Of the ten variables, eight have ratios of $(RUCC = 6, 7) / (RUCC = 4, 5) > 1$, and only one (*UNMET*) is significantly less than one. Thus, there seems to be a general pattern of the effect of *POPSIZE* increasing monotonically as one moves from $RUCC \leq 3$ to $RUCC = 4, 5$ to $RUCC = 6, 7$.

However, the monotonicity patterns ends for *RUCC*$= 8, 9$. For *RUCC*$= 8, 9$, one might argue that there is some evidence in support of

$$\left|\frac{(RUCC = 8, 9)}{(RUCC = 4, 5)}\right| \leq \left|\frac{(RUCC = 6, 7)}{(RUCC = 4, 5)}\right|$$

or

$$\left|\frac{(RUCC = 8, 9)}{(RUCC = 4, 5)}\right| \leq 1.$$

There are three critical issues with this pattern. The first is how to interpret it; we have made no progress here. The second is that it may be a composition effect due to a lack of monotonicity using *ADJAC*. In particular, from Table 5, we see that the proportion of non-adjacent counties for $RUCC = 8, 9$ is much higher than it is for $RUCC \leq 3$, $RUCC = 4, 5$, or $RUCC = 6, 7$. Thus, non-monotone effects in *rural* using *ADJAC* will lead to non-monotone effects in *rural* using *POPSIZE*.

The third is how to test it formally taking into account the randomness of our parameter estimates. We can formalize our conjecture in terms of a pair of test hypotheses. Let j index dependent variables, and define U as the set of utilization dependent variables. The our hypotheses can be stated as

$$\begin{aligned} H_0 &: \frac{(RUCC=6,7)_j}{(RUCC=4,5)_j} \leq 1, \left|\frac{(RUCC=8,9)_j}{(RUCC=4,5)_j}\right| \geq \left|\frac{(RUCC=6,7)_j}{(RUCC=4,5)_j}\right| \ \forall j \in U \qquad (1) \\ H_A &: \frac{(RUCC=6,7)_j}{(RUCC=4,5)_j} > 1, \left|\frac{(RUCC=8,9)_j}{(RUCC=4,5)_j}\right| < \left|\frac{(RUCC=6,7)_j}{(RUCC=4,5)_j}\right| \ \forall j \in U \end{aligned}$$

and

$$\begin{aligned} H_0 &: \frac{(RUCC=6,7)_j}{(RUCC=4,5)_j} \leq 1, \left|\frac{(RUCC=8,9)_j}{(RUCC=4,5)_j}\right| \geq 1 \ \forall j \in U \qquad (2) \\ H_A &: \frac{(RUCC=6,7)_j}{(RUCC=4,5)_j} > 1, \left|\frac{(RUCC=8,9)_j}{(RUCC=4,5)_j}\right| < 1 \ \forall j \in U. \end{aligned}$$

Both of these tests are multiple-argument one-sided hypotheses which require more complicated methods to test formally. We describe the details of this procedure in Appendix 9.6. The relevant test statistic for each hypthesis is

$$W^* = \min_{\theta_0 \in \Theta_0} \left\| \widehat{\theta} - \theta_0 \right\|_{\widehat{\Omega}}$$

where θ is the vector of parameters associated with the hypothesis, Θ_0 is the subset of the space for θ consistent with the null hypothesis, and $\widehat{\theta}$ is the estimate of θ, $\widehat{\Omega}$ is a consistent estimate of the covariance matrix of $\widehat{\theta}$, and $\|x\|_A$ is the quadratic form, $x'A^{-1}x$. Under conditions described in the appendix, the distribution of W^* is a mixture of χ^2 distributions. For this application, the test statistic for the hypothesis in equation (1) is 9.42 with a p-value of 0.07, and the test statistic for the hypothesis in equation (2) is 38.21 with a p-value less than 0.01. Thus, there seems to be strong evidence in support of the patterns describe in equation (2) even though we do not know how to interpret it.

5.3 County Characteristics

One might worry that the specification of *rural* might affect estimates of other variables, especially other county characteristics. Overall, we found that most of the medical health professionals variables and other medical resources variables, listed in Table 2B, had small and insignificant effects on outcomes, and their estimated effects were extremely robust across different specifications of *rural*. We wondered whether the insignificance of the county variables was caused by inclusion of site dummies. On further inspection, we saw that, with a small number of exceptions, the estimates of the county variables were not sensitive to the inclusion of site dummies. This was a bit surprising in that one would expect county characteristics to be highly correlated within a site relative to their correlation across sites. This points out the importance of recognizing that large geographic areas are often made up of several counties each of which has its own variation in socio-economic status and in health resources. This variation within large communities is challenging for health planners.

One might also wonder if there is significant variation in outcomes across sites. We can test this straightforwardly with a Wald test for H_0 : [all of the site dummy coefficients are equal] against H_A : [some of the site dummy coefficients are not equal]. The Wald statistics are overwhelmingly significant for each dependent variable. Thus, there is important variation in all of the dependent variables across sites not captured by the other explanatory variables (including the rural variables).

We also tested to see whether the rural terms interacted with race (described in Table 2C) were jointly significantly different from zero[27] and whether all of the person-county interaction terms described in Table 2C were jointly significantly different from zero.[28] We included some interactions that do not directly involve *rural* because we were concerned that they might be highly correlated with *rural*. The results of the Wald tests for these two tests are reported in Tables 8A, 8B, and 8C.[29] The race/rural interactions are generally significant for the mental health dependent variables and the health care provision variables. They have mixed results for the physical health dependent variables. The person-county interactions are significant for all three groups of dependent variables. There are three types of person-county interactions: a) between race and *rural* and b) between race and the percent of the county of the same race, and c) between race and ratio of health professionals of the same race. The Wald tests for person-county interactions suggest that having health professionals the same race as the patient might be important for outcomes. We do not know what race health professional a person of a given race chooses; we only can say that outcomes are better when the availability of matching race professionals is greater. This implies a need for further study to determine optimal ratios. The results reported in Tables 8A, 8B, and 8C show that the interaction terms are important, and the reported cell sizes in Table 5 show that using a fine disaggregation of *rural* will make it difficult to estimate interaction effects. Together, they imply that the tradeoff between inclusion of important person-county interactions and richness of the *rural* specification is binding.

[27] H_0 is that *African-American∗RUCC*$=(4,6,8)$; *African-American∗RUCC*$=(5,7,9)$; *Hispanic∗RUCC*$=(4,6,8)$; and *Hispanic∗RUCC*$=(5,7,9)$ all have zero coefficients.

[28] H_0 is that *African-American∗RUCC*$=(4,6,8)$; *African-American∗RUCC*$=(5,7,9)$; *Hispanic∗RUCC*$=(4,6,8)$; *Hispanic∗RUCC*$=(5,7,9)$; *African-American∗%African-American*; *Hispanic∗%Hispanic*; *African-American∗African-American Health Professionals/10K African-Americans*; and *Hispanic∗Hispanic Health Professionals/10K Hispanics* all have coefficients equal to zero.

[29] We report these only for the *ADJAC rural* specification because the estimates of these interaction effects are quite robust to the *rural* specification.

Table 8A: More Wald Tests for Mental Health Dependent Variables[30]		
Name	Race/Rural Interactions	All Person-County Interactions
DEPRESS	9.67**	20.68**
EMOLIM	24.60**	32.60**
MCS12	18.56**	26.81**
MENVIS	20.69**	35.13**

Table 8B: More Wald Tests for Physical Health Dependent Variables		
Name	Race/Rural Interactions	All Person-County Interactions
ARTHR	8.66*	14.77*
ASTHMA	0.38	12.60
CANCER	3.29	7.45
DIABET	3.37	16.21**
HEALIM	3.54	19.13**
HLTPRB	8.54*	34.39**
HEART	5.61	12.86
HYPTEN	7.25	13.83*
LIMIT	18.53**	25.31**
PCS12	6.32	16.93**

Table 8C: More Wald Tests for Health Care Provision Dependent Variables[31]		
Name	Race/Rural Interactions	All Person-County Interactions
DRVIST	24.96**	49.23**
ERVIST	21.68**	37.53**
HSPSTA	16.32**	42.44**
HSPEMR	10.50**	21.47**
NTHSP	279.45**	434.96**
OUTSRG	16.89**	49.29**
PUTOFF	2.20	12.28
SURGRI	17.59**	42.32**
UNMET	10.31**	15.69**

Notes:

1) Interactions for Tables 8A, 8B, and 8C are defined in Table 2C.

[30] For the race/rural interactions, the Wald statistics are distributed χ^2_4 under H_0 (with a 5% critical value of 9.49). For the person-county interactions, the Wald statistics are distributed χ^2_8 under H_0 (with a 5% critical value of 15.51).

[31] No Wald statistic is reported for *SURGOS* because the estimates of the interaction effects were estimated with no precision.

2) All tests used the *ADJAC* variable for *rural*.

Other than the site dummies, the variables that are consistently significant are the person-specific variables. The estimates associated with these variables are also robust to different specifications of *rural* and to inclusion of site dummies. Since the goal of this paper is to inform the specification of *rural*, we leave further analysis of the person-specific parameter estimates to future research.

In the process of constructing Wald tests for some of the multi-argument, one sided tests for the *POPSIZE* rural specification discussed in Section 5.2 using the methodology described in Appendix 9.6, we discovered some multicollinearity among score statistics across different dependent variables associated with utilization. Appendix 9.6 provides details about how to handle the singularity of the resulting Hessian matrix. In this section, we discuss the nature of the observed colinearity. In a model with 1100 variables,[32] there are 91 eigenvalues less than 10^{-5}. Despite the large number of problem eigenvalues, there are only a small number of patterns of variables in relevant eigenvectors associated with the problem eigenvalues. First, there are no problem eigenvalues with associated variables all from one dependent variable.[33] In other words, none of the parameter estimates were unidentified when estimating each model separately. Second, almost all of the problem eigenvalues concerned variables across the three surgery variables: *OUTSRG*, *SURGRI*, and *SURGOS*. Since $OUTSRG + SURGOS \approx SURGRI$, it should not be surprising that the inclusion of these three dependent variables causes colinearity problems. Finally, the explanatory variables associated with the problem eigenvalues are the 60 site dummies and the interactions of race and *rural*. In particular, there are only five problem eigenvalues with ten or fewer associated explanatory variables. Of these five, all involve the three surgery dependent variables. The relevant colinear explanatory variables fall into three cases: a) HMO Penetration, *Hispanic*RUCC=6,7*, and *AfrAm*RUCC=4,5*; b) *Hispanic*RUCC=6,7*, *Hispanic*%Hispanic*, and *AfrAm*RUCC=4,5*; and c) *Los Angeles*, *San Francisco*, *Hispanic*RUCC=4,5*, and *AfrAm*RUCC=6,7*. The fact that the race-*rural* interactions cause colinearity problems when using the *POPSIZE* specification is another issue one must attack when using multiple dependent variables.

6 Other Issues

One might argue that the appropriate unit of geography for community characteristics is something other than a county. For example, one might suggest using a finer measure of geography such as a zip code or census block. The advantage of using smaller units is that, to the degree that community characteristics vary

[32] There are 110 variables including 60 site dummies for each dependent variable, and there are 10 dependent variables. See Table 1C for a list of dependent variables.

[33] Throughout this discussion, when we refer to an explanatory variable associated with a problem eigenvalue, we mean the the element of the associated eigenvector corresponding to the derivative of the log likelihoods with respect to that explanatory variable is at least 0.05 in absolute value.

within counties, we would capture that variation using a smaller unit of geography. For example, Bayer, McMillan, and Rueben (2004) show that there is significant segregation across census blocks within counties. Using census tract data from 2000, we find that the mean and standard deviation of proportion African-American across US counties are respectively 0.088 and 0.145, and the mean and standard deviation of proportion African-American across US census tracts are respectively 0.138 and 0.237.[34] The mean and standard deviation are higher for census tracts than for counties because of the within-county segregation. The mean within county standard deviation is 0.051. Also some research exploits variation in distance from health care provision at a level smaller than can be captured using county data (e.g., see McClellan, McNeil, and Newhouse 1994; Newhouse and McClellan 1998; Frances et al. 2000; Athey and Stern 2002; Geweke, Gowrisankaran, and Town 2003).[35] Alternatively, one might suggest using a larger area such as a MSA or health service area (HSA),[36] a primary care services area, or a psychiatric service area. Using areas larger than counties might be more appropriate in that both health care consumers and health care providers can cross county lines to participate in the health care market.

The advantage to using county data relative to finer geographical units is that census blocks and/or zip codes are almost surely too small to describe health care markets. Zip code level data is somewhat useless in that zip code areas do not reflect communities in any real sense.[37]

While it is certainly true that county lines do not define health care markets, thus suggesting that county level data may be misleading, county level data probably still provide some important detail that cannot be captured at the HSA level. First, demographic variation within a HSA is significant and should be included (as suggested by Tables 8A, 8B, and 8C). Second, even within a HSA, geographic distance is important. For example, McClellan, McNeil, and Newhouse (1994), Newhouse and McClellan (1998), Frances et al. (2000), Athey and Stern (2002), and Geweke, Gowrisankaran, and Town (2003) use distance measures even smaller than a county, and Schootman and Shu (2004) and Litaker, Koroukian, and Love (2005) use county.

A better model of health care outcomes would measure different community characteristics at different levels of aggregation. For example, one might use racial mix variables at both the census block level and the county level and medical resources such as hospital beds, Community Mental Health Centers,

[34] We use census tracts instead of census blocks because there are many more census blocks than census tracts; so much so that information from them is difficult to collect and manage.

[35] Of those cited, all but Athey and Stern (2002) use zip codes from Medicare data to measure distance from home to various types of hospitals. Athey and Stern (2002) use some county data along with data measured at a smaller unit of geography not available in other data sources.

[36] Makuc et al (1991) defines a health service area as "one or more counties that are relatively self-contained with respect to the provision of routine hospital care."

[37] One might also worry that many variables available at the county level are not available at the census block or zip code level. However, to the degree that these variables come from underlying US Census data, they are available at the smaller level.

and HMO penetration at the HSA level. Of course, this would be feasible only if we observed residence at the census block level. In the CTS, for example, we observe the county of residence but neither the zip code nor the census block.

One way to deal with the within-county variation in some variables is to use census tract data to construct county level variables that reflect the within-county variation. For example, focusing on proportion African-American, we can construct a measure of proportion of census tract African-American for the average African-American person in the county. Let b_{ij} be the number of African-Americans in census tract j in county i and p_{ij} be the population of census tract j in county i. Define $f_{ij} = b_{ij}/p_{ij}$ as the census tract proportion (fraction) African-American, and

$$P_i = \frac{\sum_j b_{ij}}{\sum_j p_{ij}}$$

as the county proportion of African-Americans. Then

$$B_i = \frac{\sum_j b_{ij} f_{ij}}{\sum_j b_{ij}}$$

is the average proportion of neighbors a African-American person in county i has that are African-American. Similarly,

$$W_i = \frac{\sum_j (p_{ij} - b_{ij})(1 - f_{ij})}{\sum_j (p_{ij} - b_{ij})}$$

is the average proportion of neighbors a non-African-American person in county i has that are not African-American.[38] The mean values of (B_i, W_i) across counties in the US are $(0.135, 0.924)$, while the mean values of $(P_i, [1 - P_i])$ are $(0.088, 0.912)$. The correlations of (B_i, W_i) with P_i are $(0.921, -0.970)$. One can see that using the within-county variation variables has a significant effect for African-Americans and a small effect for non-African-Americans; this is because non-African-Americans represent the overwhelming majority of the population. If we run an OLS regression of B_i on a constant and P_i, we get[39]

$$\begin{aligned} B_i &= \underset{(0.002)}{0.031} + \underset{(0.009)}{1.185} P_i + e_i \\ R^2 &= 0.847. \end{aligned}$$

We also can run similar regressions separately for each state. The results are shown in Figure 4. One can see significant variation across states for

[38] This variable is available at Stern (2007).

[39] Numbers in parentheses are standard errors, and all estimates are statistically significant. We also get

$$\begin{aligned} W_i &= \underset{(0.001)}{0.999} - \underset{(0.004)}{0.854} P_i + e_i \\ R^2 &= 0.941. \end{aligned}$$

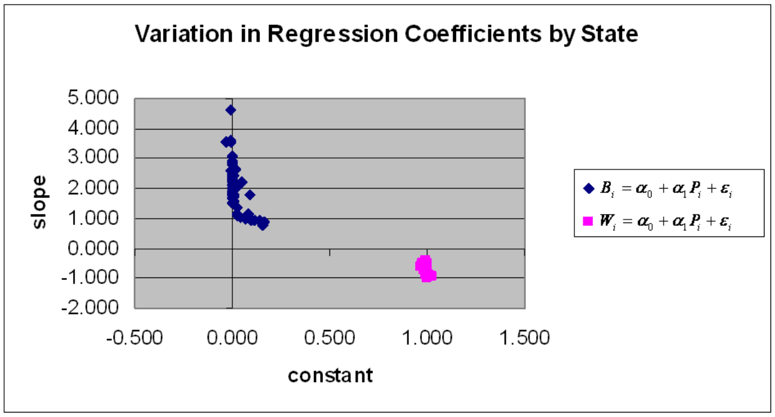

Figure 4: Variation in Regression Coefficients by State

the adjustment at the census tract level for African-American proportions and essentially no variation for nonAfrican-American proportions.

At the other extreme, though, we can measure health care provision variables at both the county level and the HSA level. There are two issues one might consider prior to using such an approach. First, even when using only county level health provision data, our estimates do not explain any significant variation in outcomes. One should not expect that adding more such variables will add much to an empirical model. Mitigating the pessimism of this argument is the idea that maybe the county health care provision variables are insignificant precisely *because* they are being measured at the wrong level of geographic aggregation; using HSA level aggregation may improve estimates and provide significant results. Second, while consumers and providers do cross county lines to participate in the health care market, distance still matters. A better approach probably would be to inversely weight health care provision effects by geographic distance (Jordan, Merwin, and Stern 2007). Such an approach allows for both cross county effects and distance effects.

Finally, one might worry that location is endogenous. It might be unreasonable to expect either consumers or health professionals to choose where to live independent of health care characteristics of the community. For example, Rosenzweig and Wolpin (1988) show that people choose when and where to move based on the availability of public programs; Moffitt (1992) surveys economic literature and finds that variation in welfare program benefits affect migration decisions; Ferreyra (2003) finds that families move to more appropriate school districts; Coffey (2003) finds that families with sensitivity to air pollution move away from locations with high pollution levels; and Vigdor (2002) provides evidence that African-Americans who live in African-American neighborhoods are selectively different than African-Americans who live in integrated neighborhoods. Walters (2002) provides an extensive survey of migration among elderly people. His results suggest that, while people in poor health might move because of poor health, it is more likely that they move closer to family. If this were the main effect of health on residential location, then it would be unlikely to cause an endogeneity problem. On the other hand, Newhouse et al (1982), Hurley (1991), Bolduc, Fortin, and Fournier (1996), and Polsky et al (2000) provide evidence that health care workers choose where to live partially as a function of the underlying health of the local population. All of these suggest that it may be inappropriate to treat location of residence as exogenous. While we have no easy solution for this problem,[40] it is the same problem for the papers, cited above, that use location as an explanatory variable or an instrument (e.g., McClellan, McNeil, and Newhouse 1994; Newhouse and McClellan 1998; Frances et al. 2000; Athey and Stern 2002; Geweke, Gowrisankaran, and Town 2003).[41]

7 Implications for Researchers

Presently, it is difficult to determine the operational definition of *rural* used in different studies. The standard for defining *rural* in research should include specifying how *rural* was defined, identifying which (if any) of existing definitions have been used. If used but modified, the modification should be stated. Sometimes the selection of which definition to use is limited by available data at particular levels of geographic analysis. This paper provides a statistical solution to choosing among rural definitions that are possible to construct based on definitions possible with data for a particular study.

A researcher could use the results of this study in two very different ways. One way would be to rely upon our empirical results to make decisions about modeling rurality. For example, if studying health care utilization, one observes from Table 6C that both the *ADJAC* and *POPSIZE* specification perform better than the *BINARY* specification and about the same as each other. However, from Table 7C, one sees that the monotonicity properties of the *ADJAC* specification are better than those of the *POPSIZE* specification. Further, the small cell sizes for rural interaction variables associated with the *POPSIZE* specification would favor the *ADJAC* specification overall. On the other hand, a study focusing on depression might be better served by the *POPSIZE* specification

[40] Moffitt (1992) also states that no literature previous to 1992 had found a good solution to the endogeneity problem.

[41] One might argue that McClellan's papers are using only differential distance between the nearest hospital and a better hospital. But this variable may vary endogenously with other unobserved health effects.

based on the results in Table 6A.

Alternatively, the researcher can use this paper as a guide to a set of tests to make an independent decision. In order to do so, one must have choices for a rural definition, one must be able to estimate separate specifications of the relevant model using the alternative specifications, and one must be able to construct Wald test statistics.[42]

8 Conclusions

This study demonstrates that the best choice for specifying *rural* within a research study varies for different dependent variables. It also shows that the effect of *rural* is frequently more subtle than can be captured by a binary rural variable such as metropolitan or non-metropolitan area, but the important variation varies with the dependent variable. Of importance is the finding that there is significant variation both within and across the sixty CTS communities.

We failed to find that there was strong evidence that *rural* affects mental or physical health problems once personal and community characteristics are included in the specification, except perhaps for depression. But *rural* consistently has a significant effect on the provision of health care services. Greater reliance on hospital care and lower reliance on outpatient care including doctor visits and emergency room visits is seen in more rural areas.

The person-county interactions were significant for all three groups of dependent variables. There are three types of person-county interactions: a) between race and *rural* and b) between race and the proportion of people in the county of the same race, and c) between race and ratio of health professionals of the same race. The Wald tests for person-county interactions suggest that having health professionals the same race as the patient might be important for outcomes. They suggest the importance of determining optimal ratios of race-specific health care providers to race-specific population and of identifying health policy initiatives to obtain optimal ratios. These results, along with the reported cell sizes in Table 5, suggest that the tradeoff between inclusion of important person-county interactions and richness of the *rural* specification is binding. The race/*rural* interactions were generally significant for the mental health dependent variables and the health care provision variables. They had mixed results for the physical health dependent variables.

9 Appendix

9.1 Rural-Urban Continuum Codes

Code Description[43]

Metro counties (based on the 2003 version of codes):

[42] Wald tests can be performed in SAS, SPSS, and Stata as options associated with statistical procedures such as regression, probit, etc.

[43] This is taken directly from US Department of Agriculture (1984).

1 Counties in metro areas of 1 million population or more
2 Counties in metro areas of 250,000 to 1 million population
3 Counties in metro areas of fewer than 250,000 population
Nonmetro counties:
4 Urban population of 20,000 or more, adjacent to a metro area
5 Urban population of 20,000 or more, not adjacent to a metro area
6 Urban population of 2,500 to 19,999, adjacent to a metro area
7 Urban population of 2,500 to 19,999, not adjacent to a metro area
8 Completely rural or less than 2,500 urban population, adjacent to a metro area
9 Completely rural or less than 2,500 urban population, not adjacent to a metro area

A nonmetro county is defined as adjacent if it physically adjoins one or more metro areas, and has at least 2% of its employed labor force commuting to central metro counties. Nonmetro counties that do not meet these criteria are classed as nonadjacent.

9.2 Other Sample Moments

This appendix includes first and second moments for all of the variables used in the study. For the explanatory variables (Tables A2 - A4), there is some small, unreported variation in moments for the different dependent variables caused by selection rules associated with dependent variables. The reported sample moments for the explanatory variables correspond to the dependent variable *HLTPRB*. Sample moments for site dummy variables are not reported.

Table A1. Dependent Variable Moments

Variable	Mean	Std Dev	Variable	Mean	Std Dev
ARTHR	0.219	0.414	*HSPSTA*	0.121	0.483
ASTHMA	0.095	0.294	*HYPTEN*	0.236	0.425
CANCER	0.044	0.204	*LIMIT*	0.225	0.418
DEPRESS	0.117	0.321	*MCS12*	52.362	9.334
DIABET	0.092	0.289	*MENVIS*	0.074	0.261
DRVIST	3.825	4.975	*NTHSP*	0.464	1.956
EMOLIM	0.121	0.327	*OUTSRG*	0.140	0.453
ERVIST	0.292	0.789	*PCS12*	48.489	10.439
HEALIM	0.095	0.294	*PUTOFF*	0.200	0.400
HEART	0.059	0.235	*SURGOS*	0.053	0.269
HLTPRB	0.744	0.437	*SURGRI*	0.193	0.533
HSPEMR	0.073	0.355	*UNMET*	0.066	0.248

Table A2. Explanatory Personal Variable Moments					
Variable	Mean	Std Dev	Variable	Mean	Std Dev
Speaks Spanish	0.036	0.186	Age	41.129	21.436
Female	0.577	0.494	Highest Grade Completed	13.620	2.596
Insurance from Private Job	0.698	0.459	Insurance from Private Company	0.099	0.299
Insurance from Other Private Source	0.023	0.149	Insurance from Medicare	0.181	0.385
Insurance from Medicaid	0.067	0.251	Insurance from Champus	0.030	0.171
Other Public Insurance	0.016	0.124	Other Insurance	0.015	0.122
Person Smokes	0.172	0.377	Works for Pay	0.510	0.500
Family Income ($100K)	0.546	0.404	Hispanic	0.087	0.282
African-American	0.117	0.321			

Table A3. Explanatory County Variable Moments		
Variable	Mean	Std Dev
Rural Code = 4,6,8	0.073	0.260
Rural Code = 5,7,9	0.032	0.176
Active MDs/10K Census	29.072	17.105
Active ObGyns/10K Census	1.359	0.722
Child Psychiatrists/10K Census	0.230	0.211
Psychiatrists/10K Census	1.429	1.277
Dos/10K Census	1.904	2.625
Nurse Practitioners/ 10K Census	2.888	2.205
CMHCs/10K Census	0.045	0.126
HMO Penetration Rate	0.313	0.169
% African-American	12.633	12.566
% Hispanic	10.890	13.877
Median Household Income	44.964	10.603
% Families in Poverty	10.583	3.892
Median Home Value	139.267	77.714
RNs/10K Census	82.491	23.635
Psychologists/10K Census	6.844	4.990
Counselors/10K Census	20.474	7.460
Social Workers/10K Census	23.092	7.902
Hospital Beds/10K Census	0.315	0.413
Hispanic Health Professionals/ 10K Hispanics	69.265	66.914
African-American Health Professionals/ 10K African-Americans	93.995	105.676

Table A4. Explanatory Personal/Community Variable Interaction Moments		
Variable	Mean	Std Dev
Hispanic * Rural Code = 4,6,8	0.002	0.043
Hispanic * Rural Code = 5,7,9	0.001	0.027
Hispanic * % Hispanic	6.104	20.524
African-American * Rural Code = 4,6,8	0.007	0.081
African-American * Rural Code = 5,7,9	0.004	0.060
African-American * % African-American	2.909	9.568
Hispanic * Hispanic Health Professionals/10K Hispanics	0.789	3.465
African-American * African-American Health Professionals/10K African-Americans	2.937	8.536

9.3 Correlated Probit

Let y_{ij}^* be the latent variable associated with family member j in family i, $j = 1, 2, .., J_i$, and $i = 1, 2, .., n$, and assume that

$$\begin{aligned} y_{ij}^* &= X_{ij}\beta + u_i + \varepsilon_{ij}, \\ u_i &\sim iidN\left(0, \sigma_u^2\right), \\ \varepsilon_{ij} &\sim iidN\left(0, \sigma_\varepsilon^2\right) \end{aligned}$$

where X_{ij} is a vector of personal characteristics specific to person j, u_i is a family-specific random effect with $u_i \sim iidN\left(0, \sigma_u^2\right)$, and ε_{ij} is a person-specific effect with $\varepsilon_{ij} \sim iidN\left(0, \sigma_\varepsilon^2\right)$. Without loss of generality and in the interest of identification, we set $\sigma_\varepsilon^2 = 1$. We define the dependent variable as

$$y_{ij} = 1 \text{ iff } y_{ij}^* > 0.$$

Then the log likelihood contribution for family i is

$$L_i = \log \int \left(\prod_{j=1}^{J_i} \Phi\left(X_{ij}\beta + u\right)^{y_{ij}} \left[1 - \Phi\left(X_{ij}\beta + u\right)\right]^{1-y_{ij}} \right) \frac{1}{\sigma_u} \phi\left(\frac{u}{\sigma_u}\right) du$$

where $\Phi(\cdot)$ is the standard normal distribution function and $\phi(\cdot)$ is the standard normal density function. It can be approximated well with K-point Gaussian quadrature (Butler and Moffitt, 1982) as

$$\log \sum_{k=1}^{K} \omega_k \left(\prod_{j=1}^{J_i} \Phi\left(X_{ij}\beta + \sigma_u\eta_k\right)^{y_{ij}} \left[1 - \Phi\left(X_{ij}\beta + \sigma_u\eta_k\right)\right]^{1-y_{ij}} \right)$$

where $\{\omega_k, \eta_k\}_{k=1}^{K}$ are the K-point Gaussian quadrature weights and locations available in Stroud and Secrest (1966). The log likelihood function is

$$L = \sum_{i=1}^{n} L_i,$$

and the vector of parameters to maximize with respect to is $\theta = (\beta, \sigma_u)$. The value of θ that maximizes L, $\widehat{\theta}$, is the maximum likelihood estimator (MLE) of θ, and it is consistent, efficient, and has an asympotic distribution of

$$\begin{aligned} \sqrt{n}\left(\widehat{\theta} - \theta\right) &\sim N(0, \Omega), \\ \Omega &= plim \left(\frac{1}{n} \sum_{i=1}^{n} \frac{\partial \log L_i}{\partial \theta} \frac{\partial \log L_i}{\partial \theta'} \right)^{-1}. \end{aligned}$$

9.4 Correlated Counts

Let y_{ij} be a count outcome variable for person j of family i. Assume that $y_{ij} \sim Poisson\left(\lambda_{ij}^{*}\right)$ with

$$\log \lambda_{ij}^{*} = X_{ij}\beta + u_i + \varepsilon_{ij}$$

where X_{ij} is a vector of personal characteristics specific to person j, u_i is a family-specific random effect with $u_i \sim iidN\left(0, \sigma_u^2\right)$, and $log\varepsilon_{ij}$ is a person-specific effect with $\varepsilon_{ij} \sim iidGamma\left(\delta, \delta\right)$. Define

$$\log \lambda_{ij}\left(u\right) = X_{ij}\beta + u$$

Then the likelihood contribution for family i is

$$\begin{aligned} L_i &= \int \prod_j \left(\frac{\lambda_{ij}\left(u\right)}{y_{ij}!}\right)\left(\frac{\delta}{\delta + \sum_j \lambda_{ij}\left(u\right)}\right)^{\delta}\left(\delta + \sum_j \lambda_{ij}\left(u\right)\right)^{-\sum_j y_{ij}} \cdot \\ &\quad \frac{\Gamma\left(\delta + \sum_j y_{ij}\right)}{\Gamma\left(\delta\right)}\frac{1}{\sigma_u}\phi\left(\frac{u}{\sigma_u}\right)du, \end{aligned}$$

and the log likelihood function is

$$L = \sum_{i=1}^{n} \log L_i.$$

Maximization of L over $\theta = \left(\beta, \delta, \sigma_u\right)$ provides consistent estimates of θ with an asymptotic distribution,

$$\begin{aligned} \sqrt{n}\left(\widehat{\theta} - \theta\right) &\sim N\left(0, V\right), \\ V &= plim\left[\frac{1}{n}\sum_{i=1}^{n}\frac{\partial \log L_i}{\partial \theta}\frac{\partial \log L_i}{\partial \theta'}\right]^{-1}. \end{aligned}$$

In the special case when $\sigma_u = 0$, the estimator becomes the Count estimator with random effects described in Hausman, Hall, and Griliches (1984). As $\delta \to \infty$ as well, the model becomes the standard Poisson count model.

9.5 Wald Tests

We are interested in two tests:

$$\begin{aligned} H_0 &: \beta_r = 0 \text{ vs. } H_A : \beta_r \neq 0; \\ H_0 &: \beta_{r1} = \beta_{r2} = .. = \beta_{rR} \text{ vs. } H_A : \beta_{r1} \neq \beta_{r2} \neq .. \neq \beta_{rR}. \end{aligned}$$

We can write both of the null hypotheses in the form

$$A\beta_r = 0$$

where β_r is the vector of rural coefficients, R is the number of rural dummies ($R = 2$ for *ADJAC*, and $R = 3$ for *POPSIZE*), $A = I$ for the first null hypothesis and

$$A = \begin{pmatrix} 1 & -1 & 0 & \cdots & 0 \\ 1 & 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & -1 \end{pmatrix}$$

for the second null hypothesis. For each of the three types of estimators used, we can represent the asymptotic distribution of the *rural* estimates $\widehat{\beta}_r$ as

$$\sqrt{n}\left(\widehat{\beta}_r - \beta_r\right) \sim N\left(0, \Omega\right).$$

This implies that, under H_0,

$$\sqrt{n}\left(A\widehat{\beta}_r - A\beta_r\right) = \sqrt{n}A\widehat{\beta}_r \sim N\left(0, A\Omega A'\right)$$

and

$$n\left(A\widehat{\beta}_r\right)'\left[A\Omega A'\right]^{-1}\left(A\widehat{\beta}_r\right) \sim \chi^2_K$$

where K is the number of restrictions implied by H_0 (i.e., the number of rows in A).

9.6 Multiple-Argument One-Sided Test Statistics

The hypotheses proposed at the end of Section 5.2 involve multiple one-sided restrictions. While the methodology for multiple restrictions is straightforward (and described in Appendix 9.5 above, and the methodology for single, one-sided restrictions is also straightforward and can be done with a t-test statistic, the methodology for multiple, one-sided restrictions is significantly more difficult. Consider a parameter vector $\theta \in \Theta$, and consider the null hypothesis $H_0 : \theta \in \Theta_0$ against $H_A : \theta \notin \Theta_0$ where $\Theta_0 \subset \Theta$ with positive measure. Now consider a Wald-type test statistic of the form $W = \left\|\widehat{\theta} - \theta_0\right\|_{\widehat{\Omega}}$ for some $\theta_0 \in \Theta_0$ where $\widehat{\theta}$ is an unrestricted consistent estimate of θ, $\widehat{\Omega}$ is a consistent estimate of the covariance matrix of $\widehat{\theta}$, and $\|x\|_A$ is the quadratic form, $x'A^{-1}x$. If we think of $H_0 : \theta = \theta_0$, then W is a Wald statistic and has an asymptotic χ^2 distribution. There are two problems with this approach. First, $H_0 : \theta = \theta_0$ is not the correct null hypothesis, and it is not obvious how to choose $\theta_0 \in \Theta_0$. Kudo (1962), Gourieroux, Holly, and Monfort (1982), and Kodde and Palm (1986) suggest defining a different test statistic,

$$W^* = \min_{\theta_0 \in \Theta_0} \left\|\widehat{\theta} - \theta_0\right\|_{\widehat{\Omega}}, \tag{3}$$

and Kodde and Palm (1986) show that

$$\lim_{N\to\infty} \Pr\left[W^*_N > c\right] = \sum_{k=0}^{K} \omega\left(K, k, \Omega\right) \Pr\left[\chi^2_k > c\right]$$

where N is the sample size, K is the number of restrictions under the null hypothesis,

$$\begin{aligned} \omega\left(K,k,\Omega\right) &= \Pr\left[R\left(v\right)=k\right], \\ v &\sim N\left(0,\Omega\right), \end{aligned}$$

$R\left(v\right)$ is the number of elements of $v>0$. While it is not feasible to analyze the distribution of W^* analytically, Wolak (1987) provides a simple method to simulate the distribution's critical values, and Stern (1995) uses it in a problem similar to the one in this paper.

The other problem involves constructing $\widehat{\Omega}$ for our application. θ involves parameters from different models (each with a different dependent variable) that were estimated independently from each other; thus we have no estimates of the covariances of estimates of parameters from separate models. We can solve this problem by computing score statistics together. Define $L_{ji\theta}$ as the score statistic for observation i and dependent variable $j=1,2,..,J$, and define $L'_{i\theta}=(L'_{1i\theta},L'_{1i\theta},...,L'_{Ji\theta})$. Then

$$\widehat{\Omega}_\theta=\left[\frac{1}{N}\sum_i L_{i\theta}L'_{i\theta}\right]^{-1} \tag{4}$$

provides a consistent estimate of the covariance matrix Ω_θ. Then we can take the submatrix of $\left(\widehat{\Omega}_\theta\right)^{-1}$ corresponding to those elements of θ associated with our test hypotheses.

When we evaluate equation (4) using our data and parameter estimates, we find that $\frac{1}{N}\sum_i L_{i\theta}L'_{i\theta}$ is singular. The nature of the singularity is discussed in more detail in Section 5.3. Here we just suggest a method to deal with the singularity. Decompose

$$\Omega_\theta=C\begin{pmatrix}\lambda & 0\\ 0 & 0\end{pmatrix}C'$$

where λ is a diagonal matrix with the nonzero eigenvalues of Ω_θ on the diagonal and C is the matrix of eigenvectors of Ω_θ. Define

$$H=C\begin{pmatrix}\lambda^{1/2}\\ 0\end{pmatrix}.$$

Then

$$HH'=C\begin{pmatrix}\lambda^{1/2}\\ 0\end{pmatrix}\begin{pmatrix}\lambda^{1/2}\\ 0\end{pmatrix}'C'=\Omega_\theta,$$

while

$$H'H=\begin{pmatrix}\lambda^{1/2}\\ 0\end{pmatrix}'C'C\begin{pmatrix}\lambda^{1/2}\\ 0\end{pmatrix}=\lambda$$

which is nonsingular. Define

$$\Omega_\theta^-=H\left(H'H\right)^{-2}H'$$

as a generalized inverse of Ω, and note that

$$\Omega_\theta \Omega_\theta^- \Omega_\theta = HH'H\left(H'H\right)^{-2} H'HH' = HH' = \Omega_\theta.$$

Thus, we can use the relevant submatrix of

$$\widehat{\Omega}_\theta^- = \widehat{H}\left(\widehat{H}'\widehat{H}\right)^{-2}\widehat{H}'$$

in the quadratic form in equation (3).

10 References

References

[1] Athey, Susan and Scott Stern (2002). "The Impact of Information Technology on Emergency Health Care Outcomes." *Rand Journal of Economics*. 33(3): 399-432.

[2] Bayer, P., R. McMillan, and K. Rueben (2004). "What Drives Racial Segregation? New Evidence Using Census Microdata." *Journal of Urban Economics*. 56(3): 514-535.

[3] Bolduc D., B. Fortin, and M. Fournier (1996). "The Effect of Incentive Policies on the Practice Location of Doctors: A Multinomial Probit Analysis." *Journal of Labor Economics.* 14(4): 703-732.

[4] Bond Huie, S., R. Hummer, and R. Rogers (2002). "Individual and Contextual Risks of Death Among Race and Ethnic Groups in the United States." *Journal of Health and Social Behavior*. 43: 359-381.

[5] Butler, M. and C. Beale (1994). *Rural-Urban Continuum Codes For Metropolitan And Nonmetropolitan Counties*. Agriculture and Rural Economy Division, Economic Research Service, United States Department of Agriculture.

[6] Butler, Richard and Robert Moffitt (1982). "A Computationally Efficient Quadrature Procedure for the One-Factor Multinomial Probit Model." *Econometrica*. 50: 761-764.

[7] Center for Studying Health System Change (2003). *Community Tracking Study Household Survey, 2000-200*3. Washington, DC.

[8] Coburn, A., A. MacKinney, T. McBride, K. Mueller, R. Slifkin, and M. Wakefield (2007). "Choosing Rural Definitions: Implications for Health Policy." *Issue Brief #2*. Rural Policy Research Institute Health Panel. http://www.rupri.org/Forms/RuralDefinitionsBrief.pdf.

[9] Coffey, Bentley (2003). “A Reexamination of Air Pollution's Effects on Infant Health: Does Mobility Matter?” Unpublished manuscript.

[10] De la Torre, A., K. Fickenscher, and H. Luft (1991). “The Zip Postal Code Difference: Methods to Improve Identification of Rural Subgroups.” *Agricultural Economics*. 5: 253-262.

[11] Diez-Roux, Ana, Sharon Merkin, Donna Arnett, Lloyd Chambless, Mark Massing, F. Javier Nieto, Paul Sorlie, Moyses Szklo, Herman Tyroler, and Robert Watson, (2001). “Neighborhood of Residence and Incidence of Coronary Disease.” *New England Journal of Medicine*. 345: 99-106.

[12] Ferreyra, Maria (2003). “Estimating the Effects of Private School Vouchers in Multi-District Economies.” Unpublished manuscript.

[13] Fortney, J., R. Owen, and J. Clothier (1999). “Impact of Travel Distance on the Disposition of Patients Presenting for Emergency Psychiatric Care.” *The Journal of Behavioral Health Services and Research.* 26(1): 104-108.

[14] Frances, C., M. Shlipak, H, Noguchi, P. Heidenreich, and M. McClellan (2000). “Does Physician Specialty Affect the Survival of Elderly Patients with Myocardial Infarction?” *Health Services Research*. 35(5 Pt 2): 1093-1116.

[15] Gessler, W., M. Hayes, A. Arcury, A. Skelly, S. Nash, and A. Soward (2004). “Use of Mapping Technology in Health Intervention Research.” *Nursing Outlook*. 52: 142-146.

[16] Geweke, J., G. Gowrisankaran, R. Town (2003). “Bayesian Inference for Hospital Quality in a Selection Model.” *Econometrica*. 71:(4) 1215-1238.

[17] Ghelfi, L. and T. Parker (1995). “A New County-Level Measure Of Urban Influence.” Staff Paper, Rural Economy Division, Economic Research Service, U.S. Department of Agriculture.

[18] Gourieroux, Christian, Alberto Holly, and Alain Monfort (1982). “Likelihood Ratio Test, Wald Test, and Kuhn-Tucker Test in Linear Models with Inequality Constraints on the Regression Parameters.” *Econometrica*. 50(1): 63-80.

[19] Hart, L., Eric Larson, and Denise Lishner (2005). “Rural Definitions for Health Policy and Research.” *American Journal of Public Health*. 95(7): 1149-1155.

[20] Hauenstein, E., S. Petterson, E. Merwin, V. Rovnyak, B. Heise, and D. Wagner (2006). “Rurality, Gender, and Mental Health Treatment.” *Family and Community Health*. 29(3): 169-185.

[21] Hausman, Jerry, Bronwyn Hall, and Zvi Griliches (1984). “Econometric Models For Count Data with an Application to the Patents R and D Relationship.” *Econometrica*. 52 (4): 909-938.

[22] Health Resources and Services Administration (2007). "Geographic Eligibility for Rural Health Grant Programs." http://ruralhealth.hrsa.gov/funding/eligibilitytestv2.asp.

[23] Hoggart, K. (1990). "Let's Do Away with Rural." *Journal of Rural Studies*. 6: 245-257.

[24] Hurley, J. (1991). "Physicians' Choices of Specialty, Location, and Mode: A Reexamination Within an Interdependent Decision Framework." *Journal of Human Resources*. 26(1): 47-71.

[25] Hynes, D., K. Koelling, K. Stroupe, N. Arnold, K. Mallin, M. Shohn, F. Weaver, L. Manheim, and L. Kok (2007). "Veterans' Access to and Use of Medicare and Veterans Affairs Health Care." *Medical Care*. 45(3): 214-223.

[26] Jordan, Lorraine, Elizabeth Merwin, and Steven Stern (2007). "The Production Function for Surgeries with Special Emphasis on Anesthesia Provision." Unpublished manuscript.

[27] Jordan, S. and D. Hargrove (1987). "Implications of an Empirical Application of Categorical Definitions of Rural." *Journal of Rural Community Psychology*. 8(2): 14-29.

[28] Kodde, David and Franz Palm (1986). "Wald Criteria for Jointly Testing Equality and Inequality Restrictions." *Econometrica*. 54(5): 1243-1248.

[29] Kudo, Akio (1962). "A Multivariate Analogue of the One-Sided Test." *Biometrika*. 50(3):403-418.

[30] Lin, S., S. Crawford, and W. Salmon (2005). "Potential Access and Revealed Access to Pain Management Medications." *Social Science and Medicine*. 60: 1881-1891.

[31] Litaker D., S. Koroukian, and T. Love (2005). "Context and Healthcare Access: Looking Beyond the Individual." *Medical Care*, 43: 531-540.

[32] Lovett, A., R. Haynes, G. Sunnenberg, and S. Gale (2002). "Car Travel Time and Accessibility by Bus to General Practioner Services: A Study Using Patient Registers and GIS." *Social Science and Medicine*. 55: 97-111.

[33] Makuc, D., B. Haglund, D. Ingram, J. Kleinman, and J. Feldman (1991). *Health Service Areas for the United States*. National Center for Health Statistics. Vital Health Statistics 2(112).

[34] McClellan, M. B. McNeil, and J. Newhouse (1994). "Does More Intensive Treatment of Acute Myocardial Infarction in the Elderly Reduce Mortality? Analysis Using Instrumental Variables." *Journal of American Medical Association*. 272(11): 859-866.

[35] Moffitt, Robert (1992). "Incentive Effects of the United-States Welfare System: - A Review." *Journal of Economic Literature*. 30(1): 1-61.

[36] Newhouse, Joseph and Mark McClellan (1998). "Econometrics in Outcomes Research: The Use of Instrumental Variables." *Annual Review of Public Health*. 19: 17-34.

[37] Newhouse, J., A. Williams, B. Bennett, and W. Schwartz (1982). "Does the Geographical Distribution of Physicians Reflect Market Failure?" *Bell Journal of Economics*. 13(2): 493-505.

[38] Niggebrugge, A., R. Haynes, A. Jones, A. Lovett, and I. Harvey (2005). "The Index of Multiple Deprivation 2000 Access Domain: A Useful Indicator for Public Health?" *Social Science and Medicine*. 60: 2743-2753.

[39] Polsky, D., P. Kletke, G. Wozniak, and J. Escarce (2000). "HMO Penetration and the Geographic Mobility of Practicing Physicians." *Journal of Health Economics*. 19(5): 793-809.

[40] Ricketts, T. and K. Johnson-Webb (1997). What is "Rural" and How to Measure "Rurality:" *A Focus on Health Care Delivery and Health Policy.* Washington, D.C. Health Resources and Services Administration: Federal Office of Rural Health Policy.

[41] Ricketts, T., K. Johnson-Webb, and P. Taylor (1998). *Definitions of Rural: A Handbook for Health Policy Makers and Researchers*. Washington, D.C.: Health Resources and Services Administration: Federal Office of Rural Health Policy.

[42] Rosenzweig, Mark and Kenneth Wolpin (1988). "Immigration Selectivity and the Effects of Public Programs." *Journal of Public Economics*. 37(3): 265-289.

[43] Rural Assistance Center (2007). http://maps.rupri.org/circ/racrural/amirural.asp.

[44] Schootman, M. and D. Shu (2004). "Small-Area Incidence Trends in Breast Cancer. *Epidem*iology. 15: 300-317.

[45] Stern, Steven (1995). "Estimating Family Long-Term Care Decisions in the Presence of Endogenous Child Characteristics." *Journal of Human Resources*. 30(3): 551-580.

[46] Stroud, A. and D. Secrest (1966). Gaussian Quadrature Formulas. Englewood Cliffs: Prentice-Hall.

[47] United States Department of Agriculture (1984). http://www.ers.usda.gov/Briefing/Rurality/RuralUrbCon/.

[48] U.S. Department of Health and Human Services, Health Resources and Services Administration, Bureau of Health Professions, National Center for Health Workforce Analysis (2004a). *User Documentation for the Area Resource File.*

[49] U.S. Department of Health and Human Services, Health Resources and Services Administration, Bureau of Health Professions, National Center for Health Workforce Analysis (2004b). *Technical Documentation with Field Numbers for the Area Resource File (ARF) 2004 Release.*

[50] Vigdor, Jacob (2002). "Locations, Outcomes, and Selective Migration." *Review of Economics and Statistics.* 84(4): 751-755.

[51] Vuong, Quang (1989). "Likelihood Ratio Tests For Model Selection and Non-Nested Hypotheses." *Econometrica.* 57 (2): 307-333.

[52] Walters, W. (2002). "Later-life Migration in the United States: A Review of Recent Research." *Journal of Planning Literature.* 17(1): 37-66.

[53] Ware, J. , M. Kosinski, and S. Keller (1994). *SF-36 Physical and Mental Health Summary Scales: A User's Manual.* Boston, MA: Health Assessment Lab.

[54] Wolak, Frank (1987). "An Exact Test for Multiple Inequality Constraints in the Linear Regression Model." *Journal of the American Statistical Association.* 82(399): 782-793.

[55] Wong, S., C. Kao, J. Crouch, and C. Korenbrot (2006). "Rural American Indian Medicaid Health Care Services Use and Health Care Costs in California." *American Journal of Public Health.* 96(2): 363-370.

[56] Yiannakoulias, N., B. Rowe, L. Svenson, D. Schopflocher, K. Kellyey, and D. Voaklander (2003). "Zones of Prevention." *Social Science and Medicine.* 57: 2065-2073.