Do Food Stamps Cause Obesity? A Generalised Bayesian Instrumental Variable Approach in the Presence of Heteroscedasticity

Matthew Salois and Kelvin Balcombe

University of Reading

8 February 2011

Online at https://mpra.ub.uni-muenchen.de/28745/
MPRA Paper No. 28745, posted 9 February 2011 19:40 UTC
Do Food Stamps Cause Obesity? A Generalised Bayesian Instrumental Variable Approach in the Presence of Heteroscedasticity

Matthew J Salois¹ and Kelvin G Balcombe
University of Reading
PO Box 237
Reading RG6 6AR
UK

Abstract
The impact of covariates on obesity in the US is investigated, with particular attention given to the role of the Supplemental Nutrition Assistance Program. The potential endogeneity of participation in SNAP is considered as a potential problem in investigating its causal influence on obesity using instrumental variable (IV) approaches. Due to the presence of heteroscedasticity in the errors, the approach for dealing with heteroscedastic errors in Geweke (1993) is extended to the Bayesian instrumental variable estimator outlined in Rossi et al. (2005). This approach leads to substantively different findings to a standard classical IV approach to correcting for heteroscedasticity. Although findings support the contention that the SNAP participation rate is associated with a greater prevalence of obesity, the evidence for this impact is substantially weakened when using the methods introduced in the paper.

Keywords. Bayesian, Food stamps, Food Insecurity, Instrumental Variable, Heteroscedasticity, Obesity.

¹To whom correspondence should be addressed. Email: m.j.salois@reading.ac.uk. Phone: +44 0118 378 7702.
1. Introduction and Motivation

The Supplemental Nutrition Assistance Program (SNAP), formerly referred to as the Food Stamp Program, is the single largest food assistance program in the United States. Although the program plays an important role in alleviating hunger and food insecurity, concern has recently refocused on the role food stamps may play in promoting obesity among the poor (Gibson 2003; Chen, Yen, and Eastwood 2005). Although different reasons have been suggested as to why food stamps might cause obesity, the main argument is that food stamp participation can lead to poor dietary choices and behaviors (Besharov 2002, 2003). The evidence regarding the impact of food stamps on obesity is mixed as some studies find a positive association between obesity and food stamp participation (Gibson 2003; Chen, Yen, and Eastwood 2005; Baum 2007; Kimbro and Rigby 2010) while others do not (Hofferth and Curtin 2005; Jones and Frongillo 2006; Kaushal 2007; Kim and Frongillo 2007; Meyerhoefer and Pylypchuk 2008).

Unobservable factors, potentially related to both food stamp participation and obesity, complicate analysis. Since participation in food stamps is a household decision affected by factors outside the obesogenic environment, an econometric model must accommodate endogenous regressors (Gibson 2003; Meyerhoefer and Pylypchuk 2008). Unobservable differences are present between households that choose to participate in the program versus household that do not. For example, nonparticipants may view food stamps as socially undesirable. Hence, the impact of food stamps on obesity is endogenous given self-selection in the participation.

This paper further investigates the role of participation in the Supplemental Nutrition Assistance Program (SNAP) on obesity using county-level cross-sectional data on obesity prevalence rates, SNAP participation rates, and other covariates for 3,051 counties across the United States. The potential endogeneity of SNAP participation is treated within a Bayesian instrumental variable (IV) approach. However, when applied to this data the models have errors with fatter tails than would be consistent with the normality assumption. Therefore, in this paper, we extend the Bayesian treatment of t-distributed errors in IV estimation. In particular, we build on the Bayesian literature for dealing with heteroscedasticity within an IV framework that ties together the Bayesian treatment of heteroscedasticity introduced by Geweke (1993) with
the Bayesian approach to IVs in Chapter 7 of Rossi et al. (2005). When using IV estimation, heteroscedasticity can exist in both the ‘structural equation’, where the covariates may include one endogenous regressor, and the equation where the endogenous regressor is regressed on instruments (which is herein referred to as the ‘instrument equation’). Our approach can be restricted so that heteroscedasticity is present in either one of the equations only.

Heteroscedasticity is treated by modelling the variance for each error using a hierarchical prior which is Gamma distributed. The parameter in this Gamma prior is also given a Gamma prior. As shown in Geweke (1993), the formulation of the error distribution as a mixture of normals is equivalent to assuming a Student-t distribution for the errors. By using this approach the errors in each or both equations have posterior t-distributions, where the degrees of freedom for the t-distribution can be estimated (or fixed) along with the specific error variances. In extending this approach to the case of IVs, the individual error precisions no longer have a posterior Gamma distribution as is the case for the single equation case. However, the posterior for the error specific precisions can be sampled by using a Gamma proposal distribution along with an acceptance probability for each precision. The remaining parameters can be sampled using the Gibbs sampler.

The estimator derived in this paper is shown to deliver substantively different results compared to existing classical and standard Bayesian IV estimates. Although the SNAP participation rate is shown to have a significant and positive relationship with the obesity rate in the standard classical IV and Bayesian IV model, this conclusion is weakened when the Bayesian IV estimator is generalised to allow for heteroscedasticity in both the structural and instrument equations. In addition, conclusions regarding other covariates change under the generalised Bayesian IV model. This article proceeds as follows. A brief review of the literature on the relationship between food stamp participation and obesity is provided in Section 2. In Section 3, a theoretical overview is given on Bayesian IV estimation in the presence of heteroscedasticity. The data and instruments are described in Section 4. The empirical results are discussed in Section 4 and Section 5 concludes.
2. Food Stamps and Obesity

Although participation in the Food Stamp Program has declined throughout the 1990s from a record high of about 28 million people in 1994 to around 18 million people in 1999 (Figlio et al. 2000), recently participation in the program has been increasingly rapidly. In 2004, 24 million people participated in the program at a cost of over $27 billion. In 2008, 28 million people participated at a cost of over $34 billion. Several hypotheses might explain why food stamps may cause weight gain and lead to obesity. The first is based on the idea that food stamp participation can promote an eating cycle of food shortages and binges. Food stamp users often spend their monthly benefits earlier in the month which can result in periods of food deprivation towards the end of the month (Wilde and Ranney 2000; Shapiro 2005). Once next month’s benefits are received, food deprivation may be followed by a tendency towards binge eating early in the next month. Although the food stamp cycle has not been directly tested, studies show that such an abrupt pattern of food deprivation and food bingeing can lead to weight gain in adults (Keys et al. 1950; Polivy et al. 1994) and children (Dietz 1995; Fisher and Birch 1999).

A second hypothesis is that food stamp participants spend more money on food than they otherwise would if they were simply given cash as benefits. There is evidence that offering a food stamp cash-out option may result in participants buying less food (Fraker et al. 1986, 1995). Consequently, food stamp benefits may result in a higher marginal propensity to consume food than simple cash (Devaney and Fraker 1989; Devaney and Moffitt 1991). Although food stamps do have the effect of reducing problems of undernutrition, the prevalence of obesity among food stamp participants has led to calls for allowing a cash-out option as a way of reducing obesity (Besharov 2002).

A third hypothesis suggests that food stamp users select foods that are dense in calories and fats and deficient in important nutrients. Higher rates of obesity have been observed among individuals with low income and low education (Drewnowski 2004). Low-income individuals typically consume diets composed largely of processed foods that are higher in calories and saturated fat (Dowler 2003) because these foods tend to be cheaper than healthier alternatives.
like fresh fruit and vegetables. Wilde and Ranney (2000) confirm that food stamp participants generally eat more foods with added sugars and more total fats. This hypothesis is connected with the notion of food insecurity (Dinour et al. 2007). Food insecure households tend to eat less fruits and vegetables and have lower intakes of key nutrients such as dietary fiber and vitamin C (Kendall et al. 1996; Rose and Oliveira 1997).

The following analysis does not seek to determine which of the hypotheses above are supported. These hypotheses are discussed because they are commonly used to support of the main contention that participation in the Food Stamp Program is causing increased levels of obesity. The following sections are aimed at putting this central contention to rigorous testing.

3. Bayesian Estimation of with Instrumental Variables in the Presence of Heteroscedasticity

In this section, the approach outlined in Rossi et al. (2005) for Bayesian instrumental variable estimation is briefly covered. The extension to the heteroscedastic case in then explained.

3.1 The Simultaneous System

The system of equations used in this paper can be expressed as (for observations \( i = 1, \ldots, n \))

\[
y_{1,i} = x'_{i} \beta + e_{1,i}
\]  

where the regressor matrix \( x'_{i} = (x'_{0,i}, y_{2,i}) \) contains exogenous variables \( x'_{0,i} \) and a potentially endogenous variable \( y_{2,i} \) that can be expressed as

\[
y_{2,i} = z'_{i} \alpha + e_{2,i}
\]  

The joint covariance matrix for the errors \( e'_{i} = (e_{1,i}, e_{2,i}) \) is independently and identically distributed (\( iid \)) normally

\[
e_{i} \sim N (0, \Sigma) \text{ where } \Sigma = \begin{pmatrix} \sigma^{2}_{1} & \rho \\ \rho & \sigma^{2}_{2} \end{pmatrix}
\]

Endogeneity occurs in the first equation with respect to \( y_{2,i} \) when \( \rho \neq 0 \). The errors are otherwise
assumed to be uncorrelated to the regressors \((x'_{0,i}, z'_i)\) in each of the equations. It follows that

\[ e_{1,i} = \frac{\rho}{\sigma^2} e_{2,i} + v_{1,i} \] (4)

and

\[ e_{2,i} = \frac{\rho}{\sigma^1} e_{1,i} + v_{2,i} \] (5)

where \(v_{1,i}\) and \(v_{2,i}\) are \(iid\) normally with mean zero and \(Cov(v_{1,i}, v_{2,i}) = 0\). Consequently, the conditional variances are

\[ Var(v_{1,i}) = \sigma^2_{1|2} = h_{1|2}^{-1} = \sigma^2_1 - \frac{\rho}{\sigma^2_2} \] (6)

and

\[ Var(v_{2,i}) = \sigma^2_{2|1} = h_{2|1}^{-1} = \sigma^2_2 - \frac{\rho}{\sigma^2_1} \] (7)

### 3.2 Bayesian Estimation under Homoscedastic Errors

The homoscedastic system of equations above can be estimated efficiently using the procedure outlined in Chapter 7 of Rossi et al. (2005). For notational convenience only the precisions are used herein (\(h_1 = \sigma^{-2}_1, h_{1|2} = \sigma^{-2}_{1|2}\), etc). Estimation requires the transformation of the equations in [1] and [2], are specified as

\[ \tilde{y}_{1,i} = \tilde{x}'_i \beta + \tilde{v}_{1,i} \] (8)

and

\[ \tilde{y}_{2,i} = \tilde{z}'_i \alpha + \tilde{v}_{2,i} \] (9)
where

\[
\begin{align*}
\tilde{y}_{1,i} &= \sqrt{h_{1|2}} (y_{1,i} - h_{2} \rho e_{2,i}) \\
\tilde{x}_i' &= \sqrt{h_{1|2}} x_i' \\
\tilde{v}_{i,1} &= \sqrt{h_{1|2}} v_{1,i} \\
\tilde{y}_{2,i} &= \sqrt{h_{2|1}} (y_{2,i} - h_{1} \rho e_{1,i}) \\
\tilde{z}_i' &= \sqrt{h_{2|1}} z_i' \\
\tilde{v}_{i,2} &= \sqrt{h_{2|1}} v_{2,i}
\end{align*}
\]

With Normal and Wishart priors $\beta \sim N(0, V_\beta)$, $\alpha \sim N(0, V_\alpha)$, and $\Sigma^{-1} \sim \text{Wishart}(S, \nu)$ the posterior distributions of $\beta$ and $\alpha$ are conditionally normally distributed (using $Y$ to denote the data)

$$
\beta|Y, \alpha, \Sigma, \sim N\left(\hat{\beta}, \hat{V}_\beta\right)
$$

where

$$
\hat{\beta} = \hat{V}_\beta \sum_{i=1}^{n} \tilde{x}_i \tilde{y}_{1,i} \text{ and } \hat{V}_\beta = \left(V_\beta^{-1} + \sum_{i=1}^{n} \tilde{x}_i \tilde{x}_i'\right)^{-1}
$$

and

$$
\alpha|Y, \beta, \Sigma, \sim N\left(\hat{\alpha}, \hat{V}_\alpha\right)
$$

where

$$
\hat{\alpha} = \hat{V}_\alpha \sum_{i=1}^{n} \tilde{z}_i \tilde{y}_{2,i} \text{ and } \hat{V}_\alpha = \left(V_\alpha^{-1} + \sum_{i=1}^{n} \tilde{z}_i \tilde{z}_i'\right)^{-1}
$$

The estimates of the errors $e_{1,i}$ and $e_{2,i}$ can then be recovered using the equations [1] and [2] and the posterior for $\Sigma^{-1}$ is:

$$
\Sigma^{-1}|Y, \beta, \alpha \sim \text{Wishart}\left(\left(\sum_{i=1}^{n} \tilde{e}_i \tilde{e}_i' + S\right), n + \nu\right).
$$

Markov Chain Monte Carlo (MCMC) estimation then proceeds by repeatedly drawing from the conditional distributions of $\beta$ and $\alpha$ given $\Sigma^{-1}$, then drawing $\Sigma^{-1}$ given $\beta$ and $\alpha$ (see Rossi et
al. (2005) for further details).

3.4 Bayesian Estimation under with Heteroscedastic Errors (in one or both equations)

The Bayes IV estimation procedure is now extended to where each of the errors is potentially heteroscedastic. Define

\[(\sqrt{\lambda_i e_{1,i}}, \sqrt{\theta_i e_{2,i}})’ \sim N(0, \Sigma)\]  
(16)

where \(\Sigma = \begin{pmatrix} h_1^{-1} & \rho \\ \rho & h_2^{-1} \end{pmatrix} \)  
(17)

where \(\lambda_i\) and \(\theta_i\) represent a specific variances for each error. Therefore, the errors can be conditioned on each other in an extension of [4] and [5] as

\[
\sqrt{\lambda_i e_{1,i}} = h_2 \rho \sqrt{\theta_i e_{2,i}} + v_{1,i} 
\]

\[
\sqrt{\theta_i e_{2,i}} = h_1 \rho \sqrt{\lambda_i e_{1,i}} + v_{2,i} 
\]

(18)  
(19)

where \(v_{1,i}\) and \(v_{2,i}\) are normally distributed with mean zero and \(\text{Cov}(v_{1,i}, v_{2,i}) = 0\) with precisions \(h_{1|2}\) and \(h_{2|1}\), respectively (as in equations [6] and [7]). By defining the normalised quantities as follows

\[
\tilde{y}_{1,i} = \left[ \sqrt{h_{1|2} \lambda_i y_{1,i}} - \sqrt{h_{1|2} h_{22} \rho \sqrt{\lambda_i e_{1,i}}} \right] 
\]

\[
\tilde{x}' = \sqrt{h_{1|2} \lambda_i x_i} 
\]

\[
\tilde{y}_{2,i} = \left[ \sqrt{h_{2|1} \theta_i y_{2,i}} - \sqrt{h_{2|1} h_{11} \rho \sqrt{\lambda_i e_{1,i}}} \right] 
\]

\[
\tilde{z}'_i = \sqrt{h_{2|1} \theta_i z_i} 
\]

\[
\tilde{y}_{2,i} = \sqrt{h_{2} \theta_i y_{2,i}} 
\]

\[
\tilde{z}'_i = \sqrt{h_{2} \theta_i z_i} 
\]
the following normalised regressions are specified

\[ \hat{y}_{1,i} = \hat{x}_i^T \beta + \hat{v}_{1,i} \quad (21) \]

\[ \hat{y}_{2,i} = \hat{z}_i' \alpha + \hat{v}_{2,i} \quad (22) \]

\[ \tilde{y}_{2,i} = \tilde{z}_i' \alpha + \tilde{v}_{2,i} \quad (23) \]

The associated errors for the equations directly above are

\[ \hat{v}_{1,i} = \sqrt{h_{1|2}} \left( \sqrt{\lambda_{i1}} e_{1,i} - h_2 \rho \sqrt{\theta_{i2}} e_{2,i} \right) \quad (24) \]

\[ \hat{v}_{2,i} = \sqrt{h_{2|1}} \left( \sqrt{\theta_{i2}} e_{2,i} - h_1 \rho \sqrt{\lambda_{i1}} e_{1,i} \right) \quad (25) \]

\[ \tilde{v}_{2,i} = \sqrt{h_2} \sqrt{\theta_{i2}} e_{2,i} \quad (26) \]

from which it can be deduced that the errors \( \hat{v}_{1,i}, \hat{v}_{2,i} \) and \( \tilde{v}_{2,i} \) are iid normally distributed with mean zero and unit variance, and \( \text{Cov}(\hat{v}_{1,i}, \hat{v}_{2,i}) = 0 \).

### 3.5 Posterior Distributions

Under the same priors for \( \beta, \alpha \) and \( \Sigma \) as the homoscedastic case, the posterior distributions for \( \beta, \alpha \) and \( \Sigma \), conditionally on \( \{\lambda_i\} \) and \( \{\theta_i\} \), are derived in a similar way to the homoscedastic case.

\[ \beta | Y, \alpha, \Sigma, \{\lambda_i\}, \{\theta_i\} \sim N \left( \hat{\beta}, \hat{V}_\beta \right) \quad (27) \]

where

\[ \hat{\beta} = \hat{V}_\beta \sum_{i=1}^n \hat{x}_i \hat{y}_{1,i} \text{ and } \hat{V}_\beta = \left( V_\beta^{-1} + \sum_{i=1}^n \hat{x}_i \hat{x}_i^T \right)^{-1} \quad (28) \]

and

\[ \alpha | Y, \beta, \Sigma, \{\lambda_i\}, \{\theta_i\} \sim N \left( \hat{\alpha}, \hat{V}_\alpha \right) \quad (29) \]

where

\[ \hat{\alpha} = \hat{V}_\alpha \sum_{i=1}^n \hat{z}_i \hat{y}_{2,i} \text{ and } \hat{V}_\alpha = \left( V_\alpha^{-1} + \sum_{i=1}^n \hat{z}_i \hat{z}_i^T \right)^{-1} \quad (30) \]
The estimates of the errors $e_{1.i}$ and $e_{2.i}$ can then be recovered using the original equations [1] and [2]. The posterior for $\Sigma^{-1}$ is:

$$\Sigma^{-1}|Y, \alpha, \beta, \Sigma, \{\theta_i\}, \{\lambda_i\} \sim \text{Wishart} \left( \left( \sum_{i=1}^{n} \hat{e}_i \hat{e}_i' + S \right), n + v \right)$$

(31)

where $\hat{e}_i' = (\sqrt{\lambda_i}e_{1,i}, \sqrt{\theta_i}e_{2,i})$. MCMC estimation then proceeds by repeatedly drawing from the conditional distributions of $\beta$, $\alpha$, and $\Sigma^{-1}$, given $\{\lambda_i\}$ and $\{\theta_i\}$. However, for the heteroscedastic case, this must be augmented with draws of $\{\lambda_i\}$ and $\{\theta_i\}$ conditionally on $Y$, $\beta$, $\alpha$, and $\Sigma^{-1}$. The next section derives these conditional distributions.

### 3.6. Conditional Posterior Distributions for $\{\lambda_i\}$ and $\{\theta_i\}$.

The priors for $\lambda_i$ and $\theta_i$ are defined to be Gamma distributed, and dependent on the parameters $\tau_\lambda$ and $\tau_\theta$. These are the priors used in Koop (2003) for the single equation case.

$$f(\lambda_i|\tau_\lambda) \propto \exp \left( \frac{-\tau_\lambda \lambda_i}{2} \right) \lambda_i^{\tau_\lambda - 1}$$

(32)

$$f(\theta_i|\tau_\theta) \propto \exp \left( \frac{-\tau_\theta \theta_i}{2} \right) \theta_i^{\tau_\theta - 1}$$

The prior is hierarchical in that $\tau_\lambda$ and $\tau_\theta$ are also assigned Gamma distributed priors, with a common hyperparameter $\tau_0$.

$$\tau_\lambda, \tau_\theta \sim G(\tau_0, 2)$$

(33)

Using $G(\tau|\mu, \nu) = \frac{1}{\Gamma(\nu)\left(\frac{2\mu}{\nu}\right)^{\nu/2}} \tau^{\nu-1} \exp \left( -\frac{\tau\mu}{2} \right)$ the distributions for the heteroscedastic parameters are (see Appendix A1)

$$f(\lambda_i|y_{1,i}, y_{2,i}, \alpha, \beta, \Sigma, \theta_i, \tau_\lambda, \tau_\theta) \propto \frac{(\tau_\lambda + 1)}{(h_{12}e_{1,i}^2 + \tau_\lambda, (\tau_\lambda + 1) \times \exp \left( h_{12}h_{22}\rho \sqrt{\theta_i} \sqrt{\lambda_1 e_{1,i} e_{2,i}} \right)$$

(34)
and

\[
f (\theta | y_1, y_2, \alpha, \beta, \Sigma, \lambda, \tau_\lambda, \tau_\theta) = G \left( \frac{(\tau_\theta + 1)}{(h_2^2 \rho^2 + h_2^2)}, \frac{(\tau_\theta + 1)}{(\tau_\theta + 1)} \right)
\times \exp \left( \frac{h_2 \rho \sqrt{\theta^2} \sqrt{\lambda_i e_1 e_2}}{\tau_\lambda} \right)
\]

(35)

3.7. Conditional Posterior Distributions for \( \tau_\lambda \) and \( \tau_\theta \).

The posterior distributions for \( \tau_\lambda, \tau_\theta \) are, conditionally on \( \{\lambda_i\} \) and \( \{\theta_i\} \), of the same form as in the single equation case, namely (see Appendix A2)

\[
f (\tau_\lambda | Y, \alpha, \beta, \Sigma, \theta, \{\lambda_i\}, \{\theta_i\}, \tau_\theta) \propto \left( \frac{\tau_\lambda}{2} \right)^{\frac{n+\tau_\lambda}{2}} \Gamma \left( \frac{\tau_\lambda}{2} \right)^{-n}
\times \exp \left( - \left( \frac{1}{\tau_0} + \frac{1}{2} \sum_{i=1}^{n} (\ln (\lambda_i^{-1}) + \lambda_i) \right) \times \tau_\lambda \right)
\]

(36)

and

\[
f (\tau_\theta | Y, \alpha, \beta, \Sigma, \{\lambda_i\}, \{\theta_i\}, \tau_\lambda) \propto \left( \frac{\tau_\theta}{2} \right)^{\frac{n+\tau_\theta}{2}} \Gamma \left( \frac{\tau_\theta}{2} \right)^{-n}
\times \exp \left( - \left( \frac{1}{\tau_0} + \frac{1}{2} \sum_{i=1}^{n} (\ln (\theta_i^{-1}) + \theta_i) \right) \times \tau_\theta \right)
\]

(37)

3.8. Estimation

The full set of posterior conditional distributions are now stated in the section above. Therefore, MCMC estimation can proceed by taking draws from

- \( \beta | Y, \alpha, \Sigma, \{\lambda_i\}, \{\theta_i\} \) using [27];
- \( \alpha | Y, \beta, \Sigma, \{\lambda_i\}, \{\theta_i\} \) using [29];
- \( \Sigma^{-1} | Y, \alpha, \beta, \Sigma, \{\theta_i\}, \{\lambda_i\} \) using [31];
- \( \{\lambda_i\} | Y, \alpha, \beta, \Sigma, \{\theta_i\}, \tau_\lambda, \tau_\theta \) using [34]
- \( \{\theta_i\} | Y, \alpha, \beta, \Sigma, \{\lambda_i\}, \tau_\lambda, \tau_\theta \) using [35];
\[ \tau_\lambda | Y, \alpha, \beta, \Sigma, \theta, \{ \lambda_i \}, \{ \theta_i \}, \tau_\theta \text{ using } [36]; \text{ and,} \]

\[ \tau_\theta | Y, \alpha, \beta, \Sigma, \theta, \{ \lambda_i \}, \{ \theta_i \}, \tau_\lambda \text{ using } [37]. \]

All of the conditional draws can be obtained using Gibbs sampling (since they have conditional posteriors of a common form) with the exceptions of \{ \lambda_i \} and \{ \theta_i \}. However, for these two sets of parameters a Gamma proposal density can be used, in which case the acceptance probability is based on the second parts of [34] and [35] for \{ \lambda_i \} and \{ \theta_i \}, respectively. For example, if a draw of \( \lambda_i^* \) is proposed using \( G \left( \frac{(\tau_\lambda + 1)}{k_{1}\sqrt{\sigma}}, \frac{(\tau_\lambda + 1)}{k_{2}} \right) \), it is accepted with probability

\[ \max \left( 1, \frac{\exp(h_{1,2}^2 \rho \sqrt{\sigma_i} \sqrt{X_i^2 e_{1,1}^2 e_{2,1}^2})}{\exp(h_{1,2}^2 \rho \sqrt{\sigma_i} \sqrt{X_i^2 e_{1,1}^2 e_{2,1}^2})} \right) \text{, where } \lambda_i \text{ is the existing draw.} \]

4. Data and Instruments

Cross-sectional data are collected from different government sources for all 3,141 U.S. counties. Only counties in the continental U.S. are included in the analysis and after dropping missing observations, the sample size is 3,051. Table 1 lists the full set of variables included in this study along with the year, geographic level, and original data source. The dependent variable, obesity prevalence rate, is the age-adjusted percentage of adults (age > 20) in a county with body mass index (BMI) greater than or equal to 30. The prevalence of obesity for each county is based on estimates from the Behavioral Risk Factor Surveillance System (BRFSS), maintained by the Centers for Disease Control and Prevention (CDC), and data from the U.S. Census Bureau’s Population Estimates Program. Two key variables are used to describe the Supplemental Nutrition Assistance Program (SNAP). The first is the ratio of the number of SNAP participants in the county to the total county population. This variable measures the county participation rate in SNAP and is taken to be the endogenous independent variable. The second is the ratio of total average monthly SNAP benefits issued to all participants in a county to the count of SNAP participants in that county. This variable measures the average monthly benefits per participant in a given year. Both SNAP variables are obtained from the Economic Research Service (ERS), U.S. Department of Agriculture (USDA).

In regards to potential instrumental variables, one possible instrument is the density of SNAP authorized stores. The availability of stores that accept SNAP benefits provides an indi-
cation of program convenience, which is likely to have an impact on the degree of participation in the program. Specifically, this variable is defined as the number of SNAP authorized stores per 1,000 persons in the county. Possible SNAP-authorized stores include supermarkets, grocery stores (small, medium, and large), convenience stores, supercenters, warehouse clubs, and specialized food stores (e.g., bakeries, produce markets, meat and seafood markets, etc.). Store data are obtained from the SNAP Benefits Redemption Division, Food and Nutrition Service (FNS), USDA. The population data are sourced from the U.S. Census Bureau. Another potential instrument is political affiliation, which has been used as a determinant of welfare program participation, and food stamp participation in particular, in other studies (Figlio et al. 2000; Ziliak et al. 2003; Baum 2007). More liberal counties may have less strict SNAP eligibility rules that can affect participation decisions. In addition, liberal counties may be more acceptable of welfare programs in general and be less likely to criticize SNAP participants or view them negatively. The percent of the 2004 presidential election votes that went to the Democratic candidate is used as an indicator of political affiliation. This measure is obtained from the U.S. Census Bureau Statistical Abstract of the United States.

Additional independent variables capture important environmental factors associated with the obesogenic environment. The socioeconomic environment (such as race, gender, and education) has been shown to be an important predictor of obesity (Robert 1999; Wang et al. 2007). Socioeconomic variables include the percent of county residents that are white and the percent that are black. An indicator for the number of males per 100 females is included to assess the gender composition of each county. Educational status is measured by the percent of county residents with a high school education and the percent with a bachelor’s degree or higher. Economic well-being is indicated by median household income (in thousands of U.S. dollars) and the percent of county residents with household income below the poverty threshold, both are sourced from the U.S. Census Bureau. Also included is the unemployment rate from the Bureau of Labor Statistics (BLS). Population density is indicated by the number of persons per square mile of land area.

The literature suggests that the local food environment is also a strong indicator of obe-
sity prevalence (Papas et al. 2007; Feng et al. 2010). County-level variables describing the local food environment can be partitioned into two general categories: the eating-out food environment and the retail food environment. The eating-out environment is indicated by the density of full-service and fast-food (or limited-service) restaurants, both defined as the number of restaurants per 1,000 persons. Full-service restaurants are establishments that provide food services to customers on the basis of a waiter/waitress service (i.e., customers are seated while ordering and being served food and then pay after eating). Fast-food restaurants are establishments that provide food services to customers on the basis that food is ordered and paid for before eating. The retail food environment is indicated by the density of supermarkets/grocery stores, and supercenters and warehouse clubs, where density is the number of outlets per 1,000 persons. Grocery stores are establishments typically referred to as supermarkets and also include small-end grocery stores that retail food as their primary business function (this included delicatessen-type outlets that satisfy this requirement). Supercenters and warehouse clubs are establishments that, in addition to retailing food and groceries, also sell merchandise including clothing, furniture, and electronics. All food environment variables are obtained from the ERS.

Other variables capture aspects of the obesogenic environment relating to physical activity options and geography, important factors associated with obesity in other studies (Poortinga 2006; Ewing et al. 2003). Physical activity availability is given by the density of recreational & fitness facilities, measured as the number of fitness and recreation centers in a county divided by the number of county residents. Fitness and recreation centers are defined as facilities primarily engaged in activities such as exercise or recreational sports activities. Also included is a dichotomous indicator if the county is a metropolitan (=1) or non-metropolitan county (=0). Under the Office of Management and Budget (OMB) classification, counties are classified as metropolitan if they are economically tied to the central counties, as measured by the share of workers commuting on a daily basis to the central counties. Counties are classified as non-metropolitan if they are outside the boundaries of metropolitan areas and have no cities with 50,000 residents or more. Both variables are also obtained from the ERS.

Finally, food insecurity has been suggested to be important in other studies on obesity,
particularly those that focus on food stamps (Frongillo et al. 1997; Townsend et al. 2001; Gibson 2003; Kim and Frongillo 2007). State-level measures of food insecurity are available from the ERS Current Population Survey Food Security Supplements. Two measures are used to indicate the prevalence of household food insecurity. The first indicates the prevalence of low or very low food insecurity and the second indicates the prevalence of very low food insecurity. Each prevalence estimate reflects the state average between 2005 through 2007. Food insecurity variables are calculated using data collected in a special supplement to the Current Population Survey conducted by the U.S. Census Bureau. Low food insecurity is defined as having multiple food access problems but few, if any, indications of reduced food intake. Very low food insecurity is defined as having reduced food intake and disrupted eating patterns due to inadequate food resources. In most cases households with very low food insecurity reported a family member going hungry at some time during the year because of insufficient income to purchase food.

5. Results and Discussion

Table 2 reports parameter estimates from two classical estimators, OLS and IV, to provide a comparison for the parameter estimates obtained using the estimator derived in Section 3. Results from the classical IV estimator in Table 2 reflect heteroscedastic-adjusted standard errors in the main structural equation but not the instrument equation. Table 4 reports parameter estimates from two Bayesian estimators. The first is the standard IV estimator as in Rossi et al. (2005), which is the Bayesian counterpart to the classical IV estimator. The second is the generalised IV estimator derived in Section 3 which extends the approach for dealing with heteroscedastic errors in Geweke (1993) to the Bayesian approach to instrumental variable estimation outlined in Rossi et al. (2005). Readers are reminded that within a Bayesian framework, it is not strictly correct to express findings in terms of classical significance. However, an analogous approach is to indicate whether the minimum of the mass below or above zero is less than a given proportion. For example, if less than 5% of the posterior mass was below zero, then this would correspond to a classical level of 10% two tailed significance. We shall continue using the term significance when dealing with the Bayesian results as well, subject to
this proviso, since it simplifies the discussion when comparing across Classical and Bayesian estimates.

5.1 Endogeneity, Overidentification, and Heteroscedasticity

First, the potential endogeneity of the SNAP participation rate variable is examined using the Hausman F-test. The calculated value of the Hausman F-statistic in Table 2 leads to rejection of the null hypothesis of exogeneity at any conventional level of significant, therefore endogeneity is an issue. The Bayesian results presented in Table 4 provides the error variance of the structural equation ($\sigma^2_1$), the covariance between the errors of the structural equation and the instrument equation ($\rho$), and the correlation between the errors. The covariance $\rho$ is positive and over twice that of its standard deviation, in both the standard Bayesian IV model in columns 2&3 (of Table 4) and in the generalised Bayesian IV model in columns 4&5. Alternatively, the correlation between the errors can be examined. Although arguably it is relatively small (0.0666 or .1775), again it is more than double its standard error meaning that the vast majority of its mass is above zero and, in this sense discussed above, is significant. The generalised Bayesian IV model derived in this paper reveals a much larger correlation between the errors of both equations than the standard Bayesian IV model. These results taken together support the hypothesis that endogeneity is present. In addition, the value of the overidentification test statistic was $<.01$ with a corresponding p-value close to unity indicating that the null hypothesis can not be rejected (i.e., the chosen instruments are exogenous).

Heteroscedasticity implies a fat-tailed error distribution, meaning we can view the problem of heteroscedasticity as an issue of t-distributed error terms. Geweke (1993) concludes that the errors from a mixture of variances (heteroscedasticity) is in fact that same as a t-distributed error problem. Thus, the issue becomes one of looking for departures from normality in the distribution of the errors. Testing for normality using the Jarque-Bera test, reported in Table 2, results in a strong rejection of the null hypothesis of normality (the critical value is 5.99 based on a 5% significance level). The problem of fat-tails in the distribution of the errors is made more clear in Table 3, which provides the percentile distribution of the errors for a normal
distribution, the main structural equation, and the instrument equation.

For example, looking at the normal error percentile of 0.10 states that for a normal distribution, 90 percent of the values occur in the central body of the distribution with 10 percent occurring in the tails. However, the tails of the error distribution for the structural equation compose about 17 percent of the total distribution. Likewise, the tails for the error distribution for the instrument equation compose about 16 percent. In general, the results in Table 3 show that there are fat tails in the error distributions of both the structural equation and the instrument equation.

In addition, turning to the Bayesian results in Table 4, the value of $\tau_\lambda$ is about 4.74 which indicates that the distribution of the errors in the main equation is not normal but rather t-distributed (a value of 25 or above is approximately normal). The same conclusion is drawn for the instrument equation. Results from the instrument equation for the generalised Bayesian IV model with heteroscedasticity are in Table 5. The value of $\tau_\theta$, which is about 3.44, indicates that the distribution of the errors in the instrument equation are also t-distributed. In summary, the SNAP participation rate variable is found to be endogenous and heteroscedasticity is uncovered in both the main equation and instrument equation. The typical IV approach, which does not permit for heteroscedasticity in the instrument equation, is not ideally suited to this data. Therefore, the generalisation developed in this paper, which allows for heteroscedasticity in one or both the structural and instrument equation, represents a more appropriate estimator for this data.

5.2 SNAP Participation Rate

Attention is now turned to the endogenous variable and comparing the parameter estimates between the classical IV estimator and the generalised Bayesian IV estimator derived with heteroscedasticity in both the structural equation and the instrument equation. Table 2 presents the parameter estimates from the classical OLS and standard IV estimator. Table 4 presents the parameters estimates from the standard Bayesian IV (columns 2&3) and the generalised Bayesian IV. Results from the classical models in Table 2 suggest that the SNAP participation
rate has a positive relationship with the obesity rate.

Accounting for endogeneity in the classical IV model does result in a slightly smaller magnitude of the SNAP participation rate variable, roughly a third smaller than in the OLS model. Although the standard error on the SNAP participation rate estimate is higher in the classical IV model, the coefficient estimate remains significant at less than the 1% level. Moving to the estimates based on the two Bayesian estimators in Table 4, the estimates in columns 2&3 of Table 4 for the standard Bayesian IV model are very similar to the results in columns 4&5 of Table 2 for the classical IV model. Both suggest a statistically significant and positive association between the SNAP participation rate and the obesity rate. The estimates in columns 4&5 of Table 4 for the generalised Bayesian IV model with heteroscedasticity in both equations suggest that extending the approach in Rossi et al. (2005) to allow for heteroscedasticity in both the main and the instrument equation has important implications. The coefficient estimate for SNAP participation rate is less than half the magnitude in the generalised IV model (columns 4&5) than in the standard Bayesian IV model (columns 2&3): 0.11 versus 0.04, respectively. Moreover, the estimate is significant at the 1% level in both the classical standard IV and the Bayesian standard IV, which falls to the 5% level in the generalised Bayesian IV model.

Even though a positive relationship is uncovered between the SNAP participation rate and the obesity rate in the preferred model (columns 4&5 of Table 4), the association is quite small when compared to other covariates (food insecurity or supercenter density, for example). In other words, the importance of the SNAP participation rate as an explanatory factor of obesity is likely to be quite negligible. Therefore, although findings support the contention that the SNAP participation rate is positively associated with the obesity rate, the evidence for this impact is substantially weakened when using the Bayesian methods introduced in the paper.

5.3 SNAP Benefits and Food Insecurity

Since the generalised Bayesian IV is the preferred model, discussion is based on results in columns 4&5 of Table 4. The parameter estimate for average monthly SNAP benefits suggests a negative association between the obesity rate and the level of SNAP benefits received. Studies
that investigate the relationship between food stamps and obesity do not typically examine participation and benefits distinctly. While one might expect the sign on the coefficient estimate for SNAP participation and SNAP benefits to be the same, the negative sign on the SNAP benefit coefficient estimate does make sense if taken in context.

SNAP participants are generally lower-income individuals and so are at risk of choosing low-cost food items that tend to be processed, high in calories, and nutrient deficient (Dowler 2003). Thus, while SNAP benefits may help participants obtain food, if the benefits are too low they may not promote healthy food choices. Higher SNAP benefit levels, however, may increase the ability of participants to afford healthier, more expensive, food options such as fresh fruits and vegetables. This logic is also congruent with the notion of the Food Stamp cycle and the pattern of food shortages and binges; higher benefits may help prevent such cycles. While the data do not permit a formal test of these conjectures, future work that investigates the relationship between food assistance programs and obesity should look at the distinct impact between participation and benefits separately.

Results on the food insecurity variables are also of interest. Food insecurity has been linked to obesity in a number of studies (Jones and Frongillo 2007; Frongillo et al. 1997; Sarlio-Lähteenkorva and Lahelma 2001; Adams et al. 2003; Vozoris and Tarasuk 2003). Although the coefficient estimate on low or very low food insecurity suggests a positive relationship with obesity, the coefficient estimate on very low food insecurity implies a negative relationship with obesity. This result is intuitive given that very low food insecurity is associated with hunger and disrupted eating patterns due to inadequate food resources. While low food insecurity does involve food access problems, it does not involve reduced food intake. Therefore, low food insecurity often leads to poor dietary choices due to budget constraints, which means more consumption of cheap, processed foods rather than healthier food options.

5.4 Other Explanatory Variables

The percent of the population white or black is positively associated with obesity prevalence, although the percent black coefficient estimate is much larger. The percent of residents with
a Bachelor’s degree or higher has a negative association with obesity, implying an inverse relationship between education and the obesity rate. The poverty rate is found to be positively related with obesity, which is consistent with Chen et al. (2010) who found that individuals living in low-income communities with income less than 200% of the federal poverty level also had higher BMIs. The positive association between poverty and obesity is also consistent with the notion that lower-income households have worse diets. Although the coefficient estimate on the unemployment rate is not statistically significant in columns 4&5 of Table 4, it is negative and significant in the other models. Therefore, not accounting for both heteroscedasticity in both the structural and instrument equation may lead to mistaken inferences regarding the relationship between obesity and unemployment.

The parameter estimate on the density of fast-food restaurants is not significant. The literature is inconclusive on the relationship between fast-food outlets and obesity. For example, using state-level data Maddock (2004) found a positive correlation between the number of fast-food outlets and the prevalence of obesity. Jeffery et al. (2006), however, found that while eating at fast-food restaurants was positively associated with obesity, the actual density of fast-food outlets was not. The estimate on the density of full-service restaurants suggests a negative association with the obesity rate, which confirms the finding in Mehta and Chang (2008). The density of full-service restaurants may indicate an eating environment with better food options or may proxy attitudes of residents with preferences for healthier foods.

The density of recreation and fitness facilities has a negative association with the obesity rate. Other studies have found that the availability of such facilities are associated with greater physical activity (Brownson 2001; Poortinga 2006) and better health (Mobley et al. 2006). Interestingly, the density of grocery stores has a positive association with obesity, although only significant at the 10% level. The positive association obtained here may be the result of combining supermarkets and small-end grocery stores in the same measure which can have opposing effects. For example, Morland and Evenson (2009) find that areas with more small grocery stores had higher rates of obesity while Morland et al. (2006) find a negative association for supermarkets and a positive association for small-end grocery stores.
Lastly, the density of supercenters and club stores is positively related to the obesity rate. Such stores heavily promote quantity discounts and bulk purchasing. Moreover, such business venues tend not to offer foods like fresh fruits and vegetables, but instead primarily sell processed foods that have longer shelf-life (Bustillos et al. 2009). Interestingly, the coefficient estimate on the density of supercenters is not significant in the other models but is significant in the generalised Bayesian IV model. In addition, the magnitude of the estimate nearly doubles in the generalised Bayesian IV model. Likewise with the coefficient estimates on the unemployment rate, mistaken inferences could be made regarding the relationship between these variables with obesity if the generalised IV estimator is not used. These results emphasize the importance of accounting for heteroscedasticity in both the structural equation and the instrument equation, not just for the endogenous variable but for the other explanatory variables also.

5.5 Determinants of SNAP Participation

The instrument equation includes the same set of covariates as the structural equation and also includes the two instruments, the density of SNAP-authorised stores and the percent of U.S. Presidential votes Democrat. Results from the instrument equation are interpreted as determinates of county rates of participation in SNAP. Although two instruments are included, only one needs to have a relationship with the SNAP participation rate in order for identification to be satisfied. The coefficient estimate on the percent U.S. Presidential votes Democrat is not statistically significant, however, a positive association is found between the density of SNAP-authorised stores and the SNAP participation rate. Of particular interest are the food insecurity variables. Although the coefficient estimate on low or very low food insecurity is not statistically significant, the estimate on very low food security suggests a positive relationship. States that have a greater prevalence of households with hunger and disrupted eating patterns have counties with higher SNAP participation rates.

The number of males per 100 females is negatively associated with the SNAP participation rate, indicating that counties with a higher male population have a lower percentage of residents enrolling in SNAP. Other studies have found that women are more likely than men to participate
in food assistance programs (Yen et al. 2008). Coefficient estimates on percent of residents that have a high school degree and Bachelor’s degree or higher are negative and significant. Gundersen and Oliveira (2001) find that being a high school graduate decreases the probability of food stamp participation. The results in Table 5 support that contention but also find that having a college degree lowers participation rates even more, thus implying that increasing levels of education are associated with lower rates of participation in SNAP. Both the poverty rate and unemployment rate are positively associated with the SNAP participation rate. Mykerezi and Mills (2010) also find a positive relationship between food stamp program participation and state unemployment levels. The results here confirm their finding and indicate that adverse economic conditions promote higher levels of participation in SNAP.

The density of supermarket/grocery stores is negatively associated, while the density of supercenters and club stores is positively associated with the SNAP participation rate. The opposite signs of these two variables is interesting since, in general, both supermarket/grocery stores and supercenter/club stores are thought to increase the availability of affordable food. According to Morland et al. (2006) supermarkets and grocery stores can improve the quality of diets, particularly in disadvantaged areas. Supercenters and warehouse club stores, however, tend to create food deserts particularly in rural areas (Blanchard and Lyson 2003). Food deserts are created when large-scale retailers draw customers from a wide geographic radius and push small-end grocers out of business, which places low-income households at a particular disadvantage of finding low-cost food. Lastly, while the coefficient estimate on the density of full-service restaurants is not significant, the density of fast-food restaurants indicates a positive association. Previous research shows a geographic correlation between low-income areas and density of fast-food restaurants (Block et al. 2004).

6. Conclusions

In recent years participation in the Food Stamps program has been increasingly rapidly. Concurrent with the growth of the food stamps program is an increasing prevalence of household food insecurity, an economic state in which households have insufficient access to healthy and
affordable food. In 2004, 11.9\% of U.S. households reported being food insecure; as of 2008, this figure increased to 14.6\% (17 million), which is the highest ever recorded (Nord et al. 2005, 2009). Although food assistance programs aim to assuage hunger and food insecurity, recent attention has turned to the contribution of food stamps to the growing problem of obesity.

This paper investigated the relationship between SNAP participation rates and the prevalence of obesity. The potential endogeneity of SNAP participation was considered within a Bayesian IV approach. While traditional treatments of IV estimation confine heteroscedasticity to one equation, the model developed in this paper extends the Bayesian treatment of heteroscedasticity to allow heteroscedasticity in the errors of both the structural and instrument equations. This generalisation allows the errors in either the structural equation, the instrument equation, or both to have posterior t-distributions.

Comparisons of the coefficients and standard errors estimates from more traditional estimators to the generalised Bayesian IV model derived in this paper revealed important differences. For OLS and IV, (both the classical and Bayesian) revealed a positive and significant relationship between SNAP participation and obesity. However, this result was weakened when using the generalised Bayesian IV model once heteroscedasticity was accounted for. The finding that SNAP was endogenous was strengthened when allowing for heteroscedasticity.
References
Besharov D. 2002. We Are Feeding the Poor as if They’re Starving. Washington Post, December 8.


Appendices

A1 Posteriors $\lambda_i$ and $\theta_i$.

In order to derive the posteriors of $\lambda_i$ and $\theta_i$ observe the fact that the joint distribution can be expressed conditionally as

$$
f (y_{1,i}, y_{2,i} | \alpha, \beta, \lambda_i, \theta_i, \Sigma, \tau_\lambda, \tau_\theta) = f (y_{1,i} | y_{2,i}, \alpha, \beta, \lambda_i, \theta_i, \tau_\lambda, \tau_\theta) A_i \times f (y_{2,i} | \alpha, \beta, \lambda_i, \theta_i, \tau_\lambda, \tau_\theta) B_i
given the change in variables formula as

$$
A_i = \left[ \left( \frac{\partial y_{1,i}}{\partial y_{1,i}} \right) f (y_{1,i} | y_{2,i}, \alpha, \beta, \lambda_i, \theta_i, \tau_\lambda, \tau_\theta) \right] (39)
$$

$$
B_i = \left[ \left( \frac{\partial y_{2,i}}{\partial y_{2,i}} \right) f (y_{2,i} | \alpha, \beta, \lambda_i, \theta_i, \tau_\lambda, \tau_\theta) \right]
$$

The posterior distributions of $\lambda_i$ and $\theta_i$ can be obtained by observing (where $\propto$ denotes a proportionality)

$$
f (\lambda_i | y_{1,i}, y_{2,i}, \alpha, \beta, \Sigma, \theta_i, \tau_\lambda, \tau_\theta) \propto A_i \times B_i \times f (\lambda_i | \tau_\lambda) (40)
$$

$$
f (\theta_i | y_{1,i}, y_{2,i}, \alpha, \beta, \Sigma, \lambda_i, \tau_\lambda, \tau_\theta, \lambda) \propto A_i \times B_i \times f (\theta_i | \tau_\theta)
$$

In order to derive the posterior distributions it is useful to observe that

$$
v_{1,i}^2 = h_{1|2} \left( \lambda_i e_{1,i}^2 + h_2^2 \rho_2^2 \theta_i e_{2,i}^2 - 2 h_2 \rho \sqrt{\theta_i} \sqrt{\lambda_i} e_{1,i} e_{2,i} \right) (41)
$$

$$
v_{2,i}^2 = h_{2|1} \left( \theta_i e_{2,i}^2 + h_1^2 \rho_1^2 \lambda_i e_{1,i}^2 - 2 h_1 \rho \sqrt{\lambda_i} e_{1,i} e_{2,i} \right)
$$

$$
v_{2,i}^2 = h_2 \theta_i e_{2,i}^2
$$

Taking each of the components and defining $\propto$ to be proportionality with respect to the quantity $q$, then the functions $A_i$ and $B_i$ observe the following

$$
A_i \lambda_i \propto \sqrt{\lambda_i} \exp \left( \frac{-v_{1,i}^2}{2} \right) \propto \sqrt{\lambda_i} \exp \left( \frac{-h_{1|2} \lambda_i e_{1,i}^2 + 2 h_{1|2} h_2 \rho \sqrt{\lambda_i} \sqrt{\theta_i} e_{1,i} e_{2,i}}{2} \right) (42)
$$

$$
A_i \theta_i \propto \exp \left( \frac{-v_{1,i}^2}{2} \right) \propto \exp \left( \frac{-h_2 \rho^2 \theta_i e_{2,i}^2 + 2 h_2 \rho \sqrt{\theta_i} \sqrt{\lambda_i} e_{1,i} e_{2,i}}{2} \right) (43)
$$

$$
B_i \lambda_i \propto \sqrt{\theta_i} \exp \left( \frac{-v_{2,i}^2}{2} \right) \propto \lambda_i (44)
$$

$$
B_i \theta_i \propto \sqrt{\theta_i} \exp \left( \frac{-v_{2,i}^2}{2} \right) = \sqrt{\theta_i} \exp \left( \frac{-h_2 \theta_i e_{2,i}^2}{2} \right) (45)
$$
The results above imply for the conditional posterior for $\lambda_i$ obeys

$$f (\lambda_i|y_{1,i}, y_{2,i}, \alpha, \beta, \Sigma, \theta_i, \tau_\lambda, \tau_\theta) \propto A_i \times B_i \times f (\lambda_i|\tau_\lambda)$$ (46)

resulting in

$$f (\lambda_i|y_{1,i}, y_{2,i}, \alpha, \beta, \Sigma, \theta_i, \tau_\lambda, \tau_\theta) \propto \exp \left(-\frac{h_{12}^2 e_{1,i}^2 + \tau_\lambda}{2} \lambda_i^{\frac{\tau_\lambda + 1}{2} - 1} \right) \times \exp \left(h_{12}^2 \rho \sqrt{\theta_i} \sqrt{\lambda_i} e_{1,i} e_{2,i} \right)$$ (47)

The results above imply for the conditional posterior for $\theta_i$ obeys

$$f (\theta_i|y_{1,i}, y_{2,i}, \alpha, \beta, \Sigma, \lambda_i, \tau_\lambda, \tau_\theta) \propto A_i \times B_i \times f (\theta_i|\tau_\theta)$$ (48)

resulting in

$$f (\theta_i|y_{1,i}, y_{2,i}, \alpha, \beta, \Sigma, \lambda_i, \tau_\lambda, \tau_\theta) \propto \sqrt{\theta_i} \exp \left(-\frac{(h_{22}^2 \rho^2 + h_{22}) e_{2,i}^2 + \tau_\theta}{2} \theta_i \right) \times \exp \left(h_{22} \rho \sqrt{\theta_i} \sqrt{\lambda_i} e_{1,i} e_{2,i} \right)$$ (49)

From [47] and [49] the results in [34] and [35] follow.

**A2 Conditional Posterior Distributions for $\tau_\lambda, \tau_\theta$.**

The posterior distributions for the degrees of freedom parameters are (defining $Y$ as the full sample data)

$$f (\tau_\lambda|Y, \alpha, \beta, \Sigma, \{\lambda_i\}, \{\theta_i\}, \tau_\theta) \propto \prod_{i=1}^{n} f (y_{1,i}|y_{2,i}, \alpha, \beta, \Sigma, \lambda_i, \theta_i, \tau_\lambda, \tau_\theta) \times f (y_{2,i}|\alpha, \beta, \Sigma, \lambda_i, \theta_i, \tau_\lambda, \tau_\theta) \times f (\lambda_i|\tau_\lambda, \tau_\theta) f (\tau_\lambda, \tau_\theta)$$ (50)

and

$$f (\tau_\theta|Y, \alpha, \beta, \Sigma, \{\lambda_i\}, \{\theta_i\}, \tau_\lambda) \propto \prod_{i=1}^{n} f (y_{1,i}|y_{2,i}, \alpha, \beta, \Sigma, \lambda_i, \theta_i, \tau_\lambda, \tau_\theta) \times f (y_{2,i}|\alpha, \beta, \Sigma, \lambda_i, \theta_i, \tau_\lambda, \tau_\theta) \times f (\theta_i|\tau_\lambda, \tau_\theta) f (\tau_\lambda, \tau_\theta)$$ (51)

The conditional distributions of $f (y_{1,i}|y_{2,i}, \alpha, \beta, \lambda_i, \theta_i, \Sigma, \tau_\lambda, \tau_\theta)$ and $f (y_{2,i}|\alpha, \beta, \Sigma, \lambda_i, \theta_i, \tau_\lambda, \tau_\theta)$ do not depend on $\tau_\lambda, \tau_\theta$. Therefore, the degrees of freedom can also be estimated by assigning a prior to $\tau_\lambda$ and $\tau_\theta$ as independent Gamma distributions as in [33] where $\tau_0$ is the prior expected value (which is set as 25 for the empirical examples). The resulting posterior for these parameter is

$$f (\tau_\lambda|Y, \alpha, \beta, \Sigma, \theta, \{\lambda_i\}, \{\theta_i\}, \tau_\theta) = \prod_{i=1}^{n} f (\lambda_i|\tau_\lambda) f (\tau_\lambda)$$ (52)
and

\[ f (\tau_\theta | Y, \alpha, \beta, \Sigma, \{\lambda_i\}, \{\theta_i\}, \tau_\lambda) = \prod_{i=1}^{n} f (\theta_i | \tau_\theta) f (\tau_\theta) \]  

(53)

Using these results, one can directly obtain [36] and [37].
Table 1. Data Description

<table>
<thead>
<tr>
<th>Variable</th>
<th>Year</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adult obesity rate</td>
<td>2007</td>
<td>CDC/BRFSS</td>
</tr>
<tr>
<td>Percentage white</td>
<td>2008</td>
<td>U.S. Census Bureau</td>
</tr>
<tr>
<td>Percentage black</td>
<td>2008</td>
<td>U.S. Census Bureau</td>
</tr>
<tr>
<td>Males per 100 females</td>
<td>2005</td>
<td>U.S. Census Bureau</td>
</tr>
<tr>
<td>High school graduate or higher</td>
<td>2000</td>
<td>U.S. Census Bureau</td>
</tr>
<tr>
<td>Bachelor’s degree or higher</td>
<td>2000</td>
<td>U.S. Census Bureau</td>
</tr>
<tr>
<td>Median household income</td>
<td>2008</td>
<td>U.S. Census Bureau</td>
</tr>
<tr>
<td>Poverty rate</td>
<td>2008</td>
<td>U.S. Census Bureau</td>
</tr>
<tr>
<td>Unemployment rate</td>
<td>2007</td>
<td>BLS</td>
</tr>
<tr>
<td>Persons per square mile</td>
<td>2006</td>
<td>U.S. Census Bureau</td>
</tr>
<tr>
<td>Metropolitan county</td>
<td>2000</td>
<td>USDA/ERS</td>
</tr>
<tr>
<td>Recreational/fitness facility density</td>
<td>2007</td>
<td>U.S. Census Bureau</td>
</tr>
<tr>
<td>Grocery/supermarket density</td>
<td>2007</td>
<td>U.S. Census Bureau</td>
</tr>
<tr>
<td>Supercenter and club store density</td>
<td>2007</td>
<td>U.S. Census Bureau</td>
</tr>
<tr>
<td>Fast-food restaurant density</td>
<td>2007</td>
<td>U.S. Census Bureau</td>
</tr>
<tr>
<td>Full-service restaurant density</td>
<td>2007</td>
<td>U.S. Census Bureau</td>
</tr>
<tr>
<td>Low or very food insecurity</td>
<td>2005-07</td>
<td>USDA/ERS</td>
</tr>
<tr>
<td>Very low food insecurity</td>
<td>2005-07</td>
<td>USDA/ERS</td>
</tr>
<tr>
<td>Average monthly SNAP benefits</td>
<td>2006</td>
<td>USDA/ERS</td>
</tr>
<tr>
<td>Ratio of SNAP participants to area population</td>
<td>2007</td>
<td>USDA/ERS</td>
</tr>
<tr>
<td>SNAP-authorised store density</td>
<td>2008</td>
<td>USDA/FNS</td>
</tr>
<tr>
<td>Percent U.S. Presidential votes Democrat</td>
<td>2004</td>
<td>U.S. Census Bureau</td>
</tr>
<tr>
<td>Variable</td>
<td>(1) OLS</td>
<td></td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>---------</td>
<td>----------------</td>
</tr>
<tr>
<td></td>
<td>Est.</td>
<td>S.E.</td>
</tr>
<tr>
<td><strong>Endogenous Variable</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ratio of SNAP participants to area population</td>
<td>0.1438</td>
<td>0.0138</td>
</tr>
<tr>
<td><strong>Other Explanatory Variables</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average monthly SNAP benefits</td>
<td>-0.0090</td>
<td>0.0013</td>
</tr>
<tr>
<td>Low or very food insecurity</td>
<td>0.4040</td>
<td>0.0529</td>
</tr>
<tr>
<td>Very low food insecurity</td>
<td>-0.8318</td>
<td>0.1333</td>
</tr>
<tr>
<td>Percentage white</td>
<td>0.0487</td>
<td>0.0035</td>
</tr>
<tr>
<td>Percentage black</td>
<td>0.1229</td>
<td>0.0042</td>
</tr>
<tr>
<td>Males per 100 females</td>
<td>-0.0107</td>
<td>0.0054</td>
</tr>
<tr>
<td>High school graduate or higher</td>
<td>0.0644</td>
<td>0.0092</td>
</tr>
<tr>
<td>Bachelor’s degree or higher</td>
<td>-0.2111</td>
<td>0.0106</td>
</tr>
<tr>
<td>Median household income</td>
<td>0.0004</td>
<td>0.0001</td>
</tr>
<tr>
<td>Poverty rate</td>
<td>0.0764</td>
<td>0.0209</td>
</tr>
<tr>
<td>Unemployment rate</td>
<td>-0.0435</td>
<td>0.0154</td>
</tr>
<tr>
<td>Persons per square mile</td>
<td>-0.0011</td>
<td>0.0003</td>
</tr>
<tr>
<td>Metropolitan county</td>
<td>0.1306</td>
<td>0.1100</td>
</tr>
<tr>
<td>Recreational/fitness facility density</td>
<td>-2.5625</td>
<td>0.4927</td>
</tr>
<tr>
<td>Grocery/supermarket density</td>
<td>0.3972</td>
<td>0.2222</td>
</tr>
<tr>
<td>Supercenter and club store density</td>
<td>2.2365</td>
<td>2.5595</td>
</tr>
<tr>
<td>Fast-food restaurant density</td>
<td>-0.2181</td>
<td>0.1587</td>
</tr>
<tr>
<td>Full-service restaurant density</td>
<td>-0.6838</td>
<td>0.0934</td>
</tr>
<tr>
<td>Constant</td>
<td>18.8534</td>
<td>1.1633</td>
</tr>
<tr>
<td><strong>R-square</strong></td>
<td>0.5843</td>
<td></td>
</tr>
<tr>
<td><strong>Standard Error</strong></td>
<td>2.3373</td>
<td></td>
</tr>
<tr>
<td><strong>F-statistic</strong></td>
<td>228.4420</td>
<td></td>
</tr>
<tr>
<td><strong>Jarque-Bera</strong></td>
<td>361.7766</td>
<td></td>
</tr>
<tr>
<td><strong>Hausman F-Statistic</strong></td>
<td>–</td>
<td>7.1200</td>
</tr>
</tbody>
</table>

*a* Significant at the 1% level (two tailed).

*b* Significant at the 5% level (two tailed).

*c* Significant at the 10% level (two tailed).
<table>
<thead>
<tr>
<th>Proportion greater than the normal ordinate</th>
<th>Normal Distribution</th>
<th>Structural Equation</th>
<th>Instrument Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.100</td>
<td>0.166</td>
<td>0.161</td>
<td></td>
</tr>
<tr>
<td>0.050</td>
<td>0.088</td>
<td>0.091</td>
<td></td>
</tr>
<tr>
<td>0.010</td>
<td>0.033</td>
<td>0.031</td>
<td></td>
</tr>
<tr>
<td>0.005</td>
<td>0.022</td>
<td>0.022</td>
<td></td>
</tr>
<tr>
<td>0.001</td>
<td>0.12</td>
<td>0.012</td>
<td></td>
</tr>
</tbody>
</table>
Table 4. Bayesian Results

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1) Standard IV</th>
<th></th>
<th>(2) Generalised IV</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Est.</td>
<td>S.D.</td>
<td>Est.</td>
<td>S.D.</td>
</tr>
<tr>
<td>Endogenous Variable</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ratio of SNAP participants to area population</td>
<td>0.1082</td>
<td>0.0193</td>
<td>0.0434</td>
<td>0.0201</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other Explanatory Variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average monthly SNAP benefits</td>
<td>-0.0090</td>
<td>0.0014</td>
<td>-0.0107</td>
<td>0.0021</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low or very food insecurity</td>
<td>0.3868</td>
<td>0.0533</td>
<td>0.3370</td>
<td>0.0523</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Very low food insecurity</td>
<td>-0.7874</td>
<td>0.1343</td>
<td>-0.4911</td>
<td>0.1323</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percentage white</td>
<td>0.0491</td>
<td>0.0036</td>
<td>0.0497</td>
<td>0.0036</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percentage black</td>
<td>0.1228</td>
<td>0.0043</td>
<td>0.1227</td>
<td>0.0043</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males per 100 females</td>
<td>-0.0162</td>
<td>0.0058</td>
<td>-0.0147</td>
<td>0.0057</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High school graduate or higher</td>
<td>0.0616</td>
<td>0.0093</td>
<td>0.0538</td>
<td>0.0090</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bachelor’s degree or higher</td>
<td>-0.2154</td>
<td>0.0107</td>
<td>-0.1910</td>
<td>0.0105</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median household income</td>
<td>0.0004</td>
<td>0.0001</td>
<td>0.0003</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poverty rate</td>
<td>0.1083</td>
<td>0.0241</td>
<td>0.1035</td>
<td>0.0243</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unemployment rate</td>
<td>-0.0356</td>
<td>0.0159</td>
<td>-0.0088</td>
<td>0.0146</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Persons per square mile</td>
<td>-0.0011</td>
<td>0.0003</td>
<td>-0.0011</td>
<td>0.0003</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Metropolitan county</td>
<td>0.1499</td>
<td>0.1097</td>
<td>0.1391</td>
<td>0.1012</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recreational/fitness facility density</td>
<td>-2.5596</td>
<td>0.4940</td>
<td>-2.6002</td>
<td>0.4726</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grocery/supermarket density</td>
<td>0.3771</td>
<td>0.2211</td>
<td>0.3722</td>
<td>0.2088</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supercenter and club store density</td>
<td>2.7134</td>
<td>2.5884</td>
<td>5.2475</td>
<td>2.3661</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fast-food restaurant density</td>
<td>0.2452</td>
<td>0.1600</td>
<td>-0.0711</td>
<td>0.1646</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full-service restaurant density</td>
<td>-0.6939</td>
<td>0.0939</td>
<td>-0.7958</td>
<td>0.1006</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>19.3613</td>
<td>1.1858</td>
<td>20.4125</td>
<td>1.1229</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t-\text{df}<em>1 (\tau</em>\lambda)</td>
<td>–</td>
<td>–</td>
<td>4.7389</td>
<td>0.4772</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variance (\sigma_1^2)</td>
<td>5.4851</td>
<td>0.1388</td>
<td>3.3375</td>
<td>0.1649</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Covariance (\rho)</td>
<td>0.3285</td>
<td>0.1243</td>
<td>0.4813</td>
<td>0.0963</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correlation</td>
<td>0.0666</td>
<td>0.0250</td>
<td>0.1775</td>
<td>0.0349</td>
</tr>
</tbody>
</table>

\textsuperscript{a} Less than 0.5% of the smaller posterior mass above or below zero.

\textsuperscript{b} Less than 2.5% of the smaller posterior mass above or below zero.

\textsuperscript{c} Less than 5% of the smaller posterior mass above or below zero.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Est.</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SNAP-authorised store density</strong></td>
<td>89.3310</td>
<td>1.6383</td>
</tr>
<tr>
<td><strong>Percent U.S. Presidential votes Democrat</strong></td>
<td>0.0025</td>
<td>0.0023</td>
</tr>
<tr>
<td><strong>Average monthly SNAP benefits</strong></td>
<td>-0.0055</td>
<td>0.0011</td>
</tr>
<tr>
<td><strong>Low or very food insecurity</strong></td>
<td>-0.0315</td>
<td>0.0409</td>
</tr>
<tr>
<td><strong>Very low food insecurity</strong></td>
<td>0.3443</td>
<td>0.1040</td>
</tr>
<tr>
<td><strong>Percentage white</strong></td>
<td>0.0179</td>
<td>0.0032</td>
</tr>
<tr>
<td><strong>Percentage black</strong></td>
<td>0.0156</td>
<td>0.0039</td>
</tr>
<tr>
<td><strong>Males per 100 females</strong></td>
<td>-0.0456</td>
<td>0.0044</td>
</tr>
<tr>
<td><strong>High school graduate or higher</strong></td>
<td>-0.0139</td>
<td>0.0078</td>
</tr>
<tr>
<td><strong>Bachelor’s degree or higher</strong></td>
<td>-0.0372</td>
<td>0.0081</td>
</tr>
<tr>
<td><strong>Median household income</strong></td>
<td>-0.0001</td>
<td>0.0001</td>
</tr>
<tr>
<td><strong>Poverty rate</strong></td>
<td>0.1728</td>
<td>0.0197</td>
</tr>
<tr>
<td><strong>Unemployment rate</strong></td>
<td>0.2334</td>
<td>0.0126</td>
</tr>
<tr>
<td><strong>Persons per square mile</strong></td>
<td>0.0001</td>
<td>0.0002</td>
</tr>
<tr>
<td><strong>Metropolitan county</strong></td>
<td>0.3953</td>
<td>0.0841</td>
</tr>
<tr>
<td><strong>Recreational/fitness facility density</strong></td>
<td>0.7482</td>
<td>0.3888</td>
</tr>
<tr>
<td><strong>Grocery/supermarket density</strong></td>
<td>-0.8784</td>
<td>0.1895</td>
</tr>
<tr>
<td><strong>Supercenter and club store density</strong></td>
<td>11.4099</td>
<td>2.3102</td>
</tr>
<tr>
<td><strong>Fast-food restaurant density</strong></td>
<td>0.2623</td>
<td>0.1299</td>
</tr>
<tr>
<td><strong>Full-service restaurant density</strong></td>
<td>-0.1112</td>
<td>0.0796</td>
</tr>
<tr>
<td><strong>Constant</strong></td>
<td>3.8002</td>
<td>0.9827</td>
</tr>
<tr>
<td><strong>t-df2</strong></td>
<td>3.4421</td>
<td>0.2848</td>
</tr>
</tbody>
</table>

\(^a\) Less than 0.5% of the smaller posterior mass above or below zero.  
\(^b\) Less than 2.5% of the smaller posterior mass above or below zero.  
\(^c\) Less than 5% of the smaller posterior mass above or below zero.