SUPPLY AND DEMAND FOR MARRIED FEMALE LABOR: RURAL AND URBAN DIFFERENCES IN THE SOUTHERN UNITED STATES

Jeffrey Alwang and Judith I. Stallmann

Abstract

This study examined the supply of and demand for married female labor in the southern United States. Special attention was given to differences in labor force participation, labor supply, and quantities of labor supplied and demanded across rural and urban areas. Once state effects were accounted for, decisions to change participation were found not to vary by urban-rural designation. Differences in demand were fully captured by an intercept shifter and the variations in hours supplied by married females between urban and rural areas. Labor supply varied greatly with the effects of key determinants (number of children, work force experience, family income) being strongly different in rural areas. Different policies are needed to promote female labor supply in rural areas as opposed to urban areas.

Key words: labor supply, labor demand, rural/ urban, simultaneous system

Potential increases in household incomes create a motivation for job creation and training programs. There is significant evidence that secondary sources of earned income are becoming more important for households everywhere. Particularly in rural areas it has been found that many families try to combat poverty and stagnating wages by entering more than one person into the labor market (Summers et al.). Female labor force participation has grown dramatically since the 1960s. Increased job opportunities are a means of raising household incomes and promoting economic development in the United States. Technical and secondary school training is often suggested as a means of raising productivity and income for second earners. The provision of child care is frequently examined as a means of promoting married female labor supply, while tax credits for child care for two worker families are a part of the current tax code.

As the number of jobs (labor demand) is increased through policy manipulations, constraints to job demand (labor supply) need to be investigated in order to ensure effective increases in household incomes. Factors affecting the labor force participation, and labor supply, of married females have been investigated in depth in numerous studies. This paper reexamines these factors or determinants for married women in the southern United States with a focus on the differences in these determinants between urban and rural areas.

Policymakers need to know whether programs which increase labor supplies (as either an intended or unintended outcome) have different impacts in rural and urban areas. Despite this need, there is little discussion in the literature of the differences in the labor supply behavior of rural and urban married women, or among any demographic groups across rural and urban designations. These differences are either assumed not to exist in studies lumping the two groups together, or assumed to lead to completely different structural regimes in studies using only urban or only rural observations.

These assumptions have potentially serious drawbacks. In the first case, if the regimes are indeed different, then misleading results will be obtained when the differences are ignored because the coefficients will represent averages of urban and rural outcomes. In the second case, if the regimes are similar, or their differences can be captured by intercept shifters, inefficiency will result from ignoring portions of the sample. This paper examines some of these possible differences and discusses their policy implications.

There is substantial evidence that the labor force behavior of married females is different in rural and urban areas. Some of these differences are related to differing demand patterns, though different propen-

---

sities to supply labor may affect observed equilibrium quantities. Nakamura et al. (1979) showed that job opportunities for female workers varied widely between urban and rural areas of Canada; Nakamura and Nakamura (1981) documented similar differences in nearly all the United States. In addition to these demand factors, differences in fixed costs associated with work, and different preferences for work in rural areas may cause quantities of labor supplied to vary systematically from urban to rural areas. Recent studies show deep differences between rural and urban labor markets and conditions (Summers et al.), but these studies have not examined the impacts of these differences on labor supply decisions.

Studies of female labor force participation have generally failed to consider the differences between rural and urban households. Blank (1988), Blundell et al., Blau and Robins, and Nakamura and Nakamura (1985) each used various models and data sets to explain female labor supply behavior. None of these accounted for systematic differences in behavior by allowing slopes to vary from rural to urban areas, nor even included intercept shifters. Nakamura et al., though documenting large differences between the behavior of urban and rural women, only included an intercept shifter in their evaluation of Canadian data. Only Blundell et al. included variables for regional differences (i.e., regions of England); however, in their study, no distinctions were made between rural and urban designations. All of the above studies used national data from representative samples which include both urban and rural observations.

This paper reexamines female labor supply and demand across urban and rural areas of the southern United States. Specific attention is paid to structural differences in labor force participation, and demand for and supply across these areas. The methodology employed permits separation of these different aspects of labor force participation and provides information on how urban areas differ from rural areas. Implications of these findings for some of the policy issues described above are discussed.

MODEL

A standard neoclassical model of married female labor supply is used (see Heckman, 1974) to express hours supplied as a function of their determinants:

$$(1) \quad h_i = b_0 + b_1 \ln (W_{fi}) + \underline{b}'_2 Z_i^* + \varepsilon_i ,$$

where h_i are (annual) hours worked by the ith married female, W_{fi} is her offered wage rate, Z_i^* is a set of variables affecting her asking wage, ε_i is a random disturbance, and b_0, b_1, $\underline{b}_2$ are the parameters of the labor supply function.[1] Z_i^* contains variables representing the number of children in the family from different age groups, the husband's earned income,[2] and the remaining household unearned income.

The offered wage (W_{fi}) for each woman is in turn determined by vectors of personal and regional economic conditions which determine the demand for a particular woman's labor services, along with total hours worked.

An approximation of this demand may be written:

$$(2) \quad \ln W_{fi} = \alpha_0 + \underline{\alpha}'_1 Z_i + \underline{\alpha}'_2 E_i + \alpha_3 h_i + u_i,$$

where Z_i is a set of human capital variables affecting the offered wage (demand factors), E_i are regional economic characteristics, and α_0, $\underline{\alpha}'_1$, $\underline{\alpha}'_2$, α'_3 are parameters to be estimated. Z_i contains variables representing the age, education, and work experience of the woman, while E_i contains S_i, a set of southern state fixed effects (dummy variables) and R_i, a dummy variable for SMSA vs. non-SMSA residence.[3]

The specification of the simultaneous system in equations (1) and (2) is consistent with numerous alternative studies. It says that the supply of women's labor is a function of her offered wage, and factors affecting her desired wage, while demand varies according to human capital characteristics and local demand shifters. Both the supply and the demand are simultaneously determined, and, hence, W_{fi} and h_i enter the right-hand-sides of equations (1) and (2), respectively. This simultaneous treatment of labor supply and demand is implicit in the numer-

---

[1] See Stern for an exposition of how this functional form can be derived from utility maximization. The $\underline{b}_2$ (underlined) notation signifies that $\underline{b}_2$ is a vector of parameters.

[2] Notice that the wife's labor supply decisions are assumed to be conditioned on the prior decision of the husband. There is a large literature on the simultaneous nature of these decisions (see e.g. Lundberg); however, the structure of the household decision process is beyond the scope of this study. Husband's Earned Income was instrumented with husband's age and age squared and the parameters for each of the models remained essentially unchanged. This result is consistent with Mroz, who also found no evidence of endogeneity of nonwife income. This leads the authors to conclude that endogeneity of this variable is not a source of bias.

[3] S_i is a vector of state fixed effects. For each observation (i), one element of the S_i vector equals one, depending on the state of residence, and all other elements are zero. The SMSA vs. non-SMSA census bureau designation is used to distinguish urban vs. rural residence, $R_i = 1$ if from non-SMSA, 0 if from SMSA.

ous studies examining labor supply behavior while "controlling for" the endogeneity of wages (see Heckman 1974; Nakamura et al., Nakamura and Nakamura 1981). The "endogeneity of wages" results because supply and demand are presumed to be simultaneously determined. The rationale behind this treatment is that in the aggregate neither wages nor hours can be treated as fixed; they each affect the other.

The relationship between this theoretical system and empirical tests of how labor supply and demand vary from rural to urban areas, and across states in the south, needs to be outlined. Tests of structural differences between urban and rural areas are straightforward. Using the r and u superscript to denote parameter estimates derived from separate estimation of equations (1) and (2) using data from only rural and urban areas, respectively, tests of $B^r = B^u$ (in equation 1) and $A^r = A^u$ (in equation 2) where $B = [b_0, \underline{b}_1, \underline{b}_2]'$ and $A = [\alpha_0, \underline{\alpha}_1, \underline{\alpha}_2, \alpha_3]'$ can be conducted. The results from these tests provide evidence about structural differences across urban and rural areas.

The theory underlying estimation of the equation system (1) and (2) can, however, be employed to distinguish more closely between those factors affecting labor supply and those affecting labor demand. First, notice that equations (1) and (2) are implicitly derived from the simultaneous system:

To achieve identification of this system, the restrictions

(3) $h_i = b_0 + b_i \ln (W_{fi}) + \underline{b}'_2 Z_i^* + \underline{b}'_3 Z_i + \underline{b}'_4 \underline{E}_i + \varepsilon_i,$

(4) $\ln (W_{fi}) = \alpha_0 + \underline{\alpha}'_{1+} Z_i + \underline{\alpha}'_2 E_i + \alpha'_3 h_i^* + \underline{\alpha}'_4 Z_i + u_i,$ and

(5) $[\underline{b}_3' = \underline{o}, \ b_4' = \underline{o}, \ \text{and} \ \underline{\alpha}_4' = \underline{o}]$

are imposed, and Z_i, E_i, and Z_i^* are used as instruments to identify the (endogenous) wages and hours. These over-identifying restrictions are derived from theory which says that Z_i and E_i, the personal and regional characteristics affecting the offered wage, only influence the hours supplied through their impacts on wages. Similarly, Z_i^* only affects actual wages through its impact on hours supplied.

The rather standard definition of Z_i, E_i, and Z_i^* (see, e.g., Nakamura et al. and Nakamura and Nakamura 1981) contains the presumption (imposed restriction) that the state-specific dummy variables (S_i) and rural residence (R_i) only affect the supply of labor through their impact on offered wages. Thus, by using the restrictions from (5), hours supplied are functions only of wages and Z_i^*, while wages are functions of hours, the state of residence, urban/rural residence and Z_i.

State effects or rural residence can, however, plausibly have independent effects on supply, and there are important policy implications for differential supply behavior between states and place of residence. Some of these implications were mentioned above. The (over-) identifying restrictions implicit in (1) and (2) need to be carefully examined and tested.

Separating the state effects (S_i) and rural residence (R_i) from E_i and considering a different regime between rural and urban residents, (3) and (4) can be rewritten:

(6) $h_i = b_0 + b_0^r R_i + b_1^u \ln (W_{fi}) + b_1^r \ln (W_{Fi}) * R_i$
$+ \underline{b}''_2 Z_i^* + \underline{b}'^r_2 Z_i^* * R_i \ \underline{b}''_3 Z_i + \underline{b}'_3 Z_i * R_i$
$+ \underline{b}''_4 S_i + \underline{b}'_4 S_i * R_i + \varepsilon_i .$

Now, the basic (over-) identifying restrictions for (6) and (7) respectively are $\underline{b}_3^u = \underline{0}$ and $\alpha_4^u = \underline{0}$.

(7) $\ln W_i = \alpha_0 + \alpha_0^r R_i + \underline{\alpha}''_1 Z_i + \underline{\alpha}'^r_1 Z_i * R_i + \underline{\alpha}''_2 S_i$
$+ \underline{\alpha}'^r_2 S_i * R_i + \alpha_3^u h_i + \alpha_3^r h_i * R_i + \underline{\alpha}''_4 Z_i^*$
$+ \underline{\alpha}'^r_4 Z_i^* * R_i + u$

These restrictions are imposed from the start.[4] To examine whether there are differences in supplies not just attributable to demand differences (i.e., wages offered) between states and from rural to urban areas, the restrictions $b_4^u = 0$, $b_0^r = 0$, $b_2^r = 0$, $b_3^r = 0$, $b_4^r = 0$, are jointly tested. Rejection of this test of over-identifying restrictions will provide evidence that labor supply parameters differ between rural and urban areas and between states. Acceptance of the restrictions implies that urban/rural differences, if they exist, only affect labor supply through their effect on demand (wages). Depending on the results of these tests, further analyses of the sources of supply differences (if the restriction is rejected) can be conducted.

---

[4] Because Z_i and Z_i^* are <u>vectors</u> and only 1 exogenous variable needs to be deleted from each equation to achieve identification, the restrictions $\underline{b}_3 u = \underline{0}$, $\alpha_4 u = \underline{0}$ are over-identifying. These over-identifying restrictions will not be tested, because exclusion of any one of these vector components is not logical; i.e., they should be excluded together. The restrictions $\underline{b}_3 r = \underline{0}$, $\alpha_4 r = \underline{0}$ can also be imposed at this point; however, they are instead tested. The decision to impose $\underline{b}_3 u = \underline{0}$, $\alpha_4 u = \underline{0}$ is arbitrary, but some restriction is necessary to achieve identification.

The demand (wage) side of the model can be examined in a similar fashion. Tests of $\alpha_0^5 = 0$, $\underline{\alpha_1^5} = \underline{0}$, $\underline{\alpha_2^5} = \underline{0}$ and $\underline{\alpha_4^5} = \underline{0}$ will indicate if wages vary systematically from urban to rural areas, or whether any differences are solely attributable to differences in supply propensities (h_i).

DATA

The data used for this study were taken from the March 1979 Current Population Survey (CPS). In the CPS, families were asked a series of questions about their labor market experience in the previous week, and for information on household demographics. In addition to this, in the March supplement to the CPS (known as the Annual Demographic File), respondents were asked about labor force participation in the previous year, husband's earnings and family unearned income. From the national data, 33,854 single-family households were identified with husbands and wives both present. Of the national total, 6,277 were from the Southern states;[5] 3,032 of these households lived within SMSAs, and 3,245 were from outside of SMSAs. There were no observations for SMSA residents from West Virginia, South Carolina, Mississippi and Arkansas, so these states were dropped from the analysis,[6] leaving 5,326 observations, 3,032 within SMSAs and 2,294 from outside. The CPS distinction between SMSA and non-SMSA was used to define urban/rural dwellers. A description of the variables included in the analysis and their summary statistics are found in Table 1. The variable definitions directly coincide with CPS definitions.

ESTIMATION

Before the results are discussed, a feature of the empirical analysis must be introduced. Since the offered wage (W_{fi}) is only observed for those women who worked, equations (6) and (7) have to be tested and corrected for selectivity bias (see Maddala for a summary treatment of these problems). Heckman's 2-stage procedure (Heckman 1974, 1976, 1979) for checking and then correcting for selectivity was used. To check for selection bias, a probit equation was estimated for the probability of participation. The inverse mills ratio is calculated using $\lambda_i = \emptyset(Z_i, Z_i^*)/\Phi(Z_i, Z_i^*)$ where $\emptyset$ is the density and Φ the distribution of the probit index. The computed λ_i is then entered into the regression equations for the

reduced forms of (1) and (2) and the standard errors of the estimates are adjusted accordingly. The full selectivity model, as well as the nature of the corrections to the standard errors needed are presented in pp. 231-276 of Maddala.

The labor supply model contains three basic equations: a participation equation, a wage equation, and an hours-supplied equation. They are discussed in logical order.

Participation Equation

The probit female labor force participation equation was estimated separately for the SMSA/non-SMSA groups. This model (full model) includes all the variables described in Table 1. In order to account for the regional factors found in E_i, the model includes fixed state effects. This precludes the ability to test the impact of unemployment, wages, and other variables usually measured at the state level. However, the influence of these variables is fully absorbed by the fixed state effect.

The results from the participation model are presented in Table 2. The models fit the data very well as evidenced by the rather high percentage of correct predictions. Each regression is highly significant as shown by the likelihood ratio test.

Restrictions were imposed to restrict the SMSA/non-SMSA differences to an intercept shifter (restricted model), i.e., a dummy variable ($R_i = 1$ if non-SMSA, 0 otherwise). Likelihood ratio tests were performed to test the full versus the restricted models. The results of the likelihood ratio tests of equality of slope coefficients in the participation equation from rural to urban areas show that the hypothesis that the slope coefficients are equal cannot be rejected (the test statistic, with 21 degrees of freedom is $X^2(21) = 31.4$, while the critical value at a 5 percent level of confidence is X^2c (21), .95 = 32.67). The differences, if any, between the parameters of the labor force participation equation between urban and rural areas are adequately captured by an intercept shifter. As can be seen in the first column of Table 2, the non-SMSA dummy variable is not significant in the restricted model; thus, there is little strong evidence of different female labor force participation decisions from SMSA to non-SMSA areas. It should be remembered, however, that the likelihood ratio test reported above came close to rejecting the equality of slopes across rural and urban areas. In summary, differences in parameters across rural and urban areas were tested by comparing the

[5] These states are: Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, South Carolina, Virginia, West Virginia.

[6] These states cannot be included since the variables S_i and $S_i * R_i$ will be perfectly collinear for these four states.

Table 1. Description of Variables and Summary Statistics

Variable	Description	Non-SMSA ($R_i = 1$) (N = 2,294)		SMSA ($R_i = 0$) (N = 3,032)	
		Mean	S.D.	Mean	S.D.
Employment Satus	Dummy = 1 if employed	.45	.50	.49	.50
Hours	Annualized hours worked[c]	1,873.30	545.85	1,868.50	569.86
Wage	Hourly equivalent wage	4.27	5.04	5.06	3.86
Age[a]	Age of female (in years)	44.35	15.92	43.59	15.18
Highest grade[a]	Highest grade of education completed	12.25	2.94	13.33	2.75
Highest grade squared[a]		158.80	68.59	185.11	71.58
Husband's highest grade[a]		12.00	3.73	13.54	3.50
Race[a]	Dummy = 1 if non-white	.10	.30	.16	.33
Lagged participation	Dummy = 1 if female worked in prior year	.50	.50	.52	.50
Child < 1[b]	Number of children less than 1 year old in family	.05	.22	.05	.21
Child 2-6[b]	Children from 2-6 years	.23	.54	.23	.52
Child 7-12[b]	Children from 7-12 years	.35	.66	.30	.63
Child 13-18[b]	Children between 13 and 18 years	.36	.72	.34	.71
Husband's earnings[b]	Wage earnings of husband, annualized	10,906.00	9,752.10	14,900.00	12,046.00
Unearned Income[b]	Annualized unearned income of family	2,939.90	5,100.40	3,318.10	6,443.00

[a]Included in Z_i.
[b]Included in Z_i^*.
[c]Computed using "usually weekly hours on main job" question in CPS, Annualized by multiplying variable by 50.

likelihood ratio of the full versus the restricted models. The equality of slope coefficients from rural to urban areas could not be rejected. The t-test on the restricted model indicates that the intercept shifter for SMSA vs non-SMSA is not significant.

In order to examine the effect that the inclusion of the lagged participation variable had on these results, the model was reestimated while excluding this variable. The results with lagged participation variable deleted are shown in Table 3. In this model, the hypothesis that the urban-rural participation structure is equal is strongly rejected (X^2 (20) = 52.4; X^2_c (20), .95 = 31.41). This provides evidence that there are strong persistent effects in the female labor force participation decision (individual heterogeneity); when these effects are accounted for (by including lagged participation) there are only a few differences between rural and urban areas in the decision to participate by married women. When these persistent effects are not accounted for, then large strong differences in participation across rural and urban areas exist. Entering lagged participation in the re-

gression is one means of accounting for dynamic life-cycle effects (Nakamura and Nakamura 1985). Because of this, the lagged participation variable is included in the participation model used to check for selectivity in the wage and hours equation.

The results of the participation model in Tables 2 and 3 are generally consistent with prior findings and with expectations. More education raises the probability of working. The education effect is concave, though the marginal effect is positive throughout the relevant range of the variable. This concave effect reinforces the notion that non-linear effects are needed to effectively model the impact of education on labor force decisions. Few other studies examine the non-linear effects of education on participation and use instead a linear term, which is generally found to be positive (Nakamura and Nakamura 1981; Nakamura et al. 1979). Age effects are similarly concave. Age has a positive impact on female labor force participation until 33.6 years, when the likelihood of labor force participation begins declining. This is largely consistent with prior research

Table 2. Probit results for labor force participation[a]

Explanatory Variable	SMSA / Non-SMSA Constrained Equal (Restricted Model)	SMSA / Non-SMSA Allowed to Vary (Full Model)	
		SMSA	Non-SMSA
Intercept	-3.1771 (-6.66)	-1.6464 (-2.42)	-4.8025 (-6.71)
Child < 1	-.5570 (-5.25)	-0.6306 (-4.33)	-.4721 (-2.99)
Child 2-6	-.1992 (-3.94)	-0.1191 (-1.75)	-.2869 (-3.72)
Child 7-12	-.0836 (-2.12)	-.0963 (-1.76)	-.0759 (-1.31)
Child 13-18	-.1054 (-2.87)	-.0687 (-1.36)	-.1368 (-2.51)
Husband's Earnings[b]	-.0421 (-1.51)	-.0215 (-.36)	-.1039 (-2.10)
Unearned Income[b]	-.1018 (-1.93)	-.0316 (-.51)	-.3524 (-2.81)
Age	.0599 (4.31)	.0440 (2.40)	.0800 (3.68)
Age Squared	-.00089 (-5.48)	-.00073 (-3.43)	-.0011 (-4.19)
Highest Grade	.1547 (2.71)	.0180 (-.22)	.3224 (3.68)
Highest Grade Squared	.00438 (-2.01)	-.00204 (.66)	-.0112 (-3.22)
Lagged participation	2.4799 (46.98)	2.5632 (35.27)	2.4035 (28.87)
Race	.3175 (4.10)	.3866 (3.93)	.1674 (1.29)
Husband's Highest Grade	-.00761 (-.73)	-.0190 (-1.32)	.00741 (.48)
R_i	-.0822 (-1.51)	- -	- -
State Effects			
Maryland	.0243 (.21)	-.0203 (-.15)	.0244 (.10)
Virginia	-.0992 (-.93)	-.2096 (-1.56)	.1995 (1.06)
North Carolina	.1172 (1.08)	.0718 (.46)	.2243 (1.47)
Georgia	.0754 (.70)	.0612 (0.39)	.1850 (1.21)
Florida	.1393 (1.44)	.0678 (0.54)	.2800 (1.77)
Kentucky	-.0995 (-.87)	-.1259 (-.72)	-.0149 (-.10)
Tennessee	.1349 (1.17)	.1167 (.72)	.2193 (1.27)
Alabama	-.2073 (-1.84)	-.3173 (-2.13)	-.0324 (-1.19)
N	5,236	5,236	
LR[c]	4,277.5	4,246.2	
(p value)	.000	.000	
Percent correct predictions[d]	.902	.903	

[a]Asymptotic t-statistics in parentheses. Deleted state fixed effect is Louisiana.
[b]In $10.000.
[c]This likelihood ratio tests indicates the overall significance of the regression. It is analogous to the overall F-test for a linear regression.
[d]Total correct predictions (i.e., cases where $\hat{P}_i$, P_i = 1, plus cases where $\hat{P}_i$ = 0, P_i = 0, where $\hat{P}_i$ is the predicted participation for the ith person and P_i is actual participation) divided by the total number of observations.

Table 3. Probit Results for Labor Force Participation with Lagged Participation Excluded[a]

Explanatory Variable	SMSA / Non-SMSA Constrained Equal	SMSA / Non-SMSA Allowed to Vary	
		SMSA	Non-SMSA
Intercept	-1.4961	-.2702	-2.727
	(-4.23)	(-0.54)	(-5.14)
Child < 1	-.6501	-0.7156	-0.5340
	(-7.33)	(-5.94)	(-4.03)
Child 2-6	-.5786	-0.5906	-0.5677
	(-14.43)	(-11.04)	(-9.23)
Child 7-12	-.1950	-0.2035	0.1850
	(-6.43)	(-4.94)	(-4.08)
Child 13-18	-.1443	-0.1591	-0.11670
	(-5.16)	(-4.29)	(-2.71)
Husband's Earnings[b]	-.1387	-0.1197	-0.1903
	(-6.56)	(-4.61)	(-4.95)
Unearned Income[b]	-.3408	-0.2299	-0.67612
	(-8.04)	(-4.64)	(-7.16)
Age	.0784	.0647	.0967
	(7.48)	(4.69)	(5.92)
Age Squared	-.0001	-.0012	-.0014
	(-10.99)	(-7.65)	(-7.75)
Highest Grade	.0814	.00004	0.1295
	(1.94)	(0.001)	(1.99)
Highest Grade Squared	.0008	.0041	-.0013
	(0.52)	(1.86)	(-0.49)
Race	.4106	0.4014	0.3588
	(6.89)	(5.45)	(3.46)
Husband's Highest Grade	-0.0035	-.0276	.0270
	(-.45)	(-2.60)	(2.28)
Non-SMSA	-.0278	-	-
	(-.67)	-	-
Maryland	.1190	.0217	0.2094
	(1.38)	(0.21)	(1.03)
Virginia	.1352	-.0072	0.4270
	(1.65)	(-0.07)	(2.89)
North Carolina	.3554	0.1181	0.6440
	(4.34)	(1.01)	(5.32)
Georgia	.1818	.0666	0.3919
	(2.23)	(0.57)	(3.270)
Florida	.1599	.0196	0.4104
	(2.18)	(0.21)	(3.37)
Kentucky	-.0052	-0.1978	0.2428
	(-.06)	(-1.52)	(1.93)
Tennessee	.2064	.0868	0.4218
	(2.36)	(0.73)	(3.17)
Alabama	-.0019	-.0779	0.1446
	(-.02)	(-0.68)	(1.07)
N	5,231	5,236	
LR[c]	1,473.8	1,526.3	
(p value)	.000	.000	
Percent correct predictions[d]	.704	.710	

[a]Asymptotic t-statistics in parentheses. Deleted state fixed effect is Louisiana.
[b]In $10,000
[c]This likelihood ratio tests indicates the overall significance of the regression. It is analogous to the overall F-test for a linear regression.
[d]Total correct predictions (i.e., cases where $\hat{P}_i$, P_i = 1, plus cases where $\hat{P}_i$ = 0, P_i = 0, where $\hat{P}_i$ is the predicted participation for the ith person and P_i in the actual participation) divided by total number of observations.

(Nakamura and Nakamura 1981; Nakamura et al. 1979).

The effects of the presence of children on labor force participation are extremely strong and consistent with previously published results (Huffman and Lang; Nakamura and Nakamura 1981). The presence of children aged less than 6 years had a strong negative effect on labor force participation by married females. Children between 7 and 12 years of age had a smaller, but still negative effect.

The effects of the husband's earned and family's unearned income on female labor force participation are negative yet not statistically significant. This might be explained by an early decision in the lifecycle by the women to participate or not; continued participation by her is unaffected by small changes in either the husband's wage or family unearned income. When the variable representing lagged participation is deleted both these variables become negative and strongly significant (see Table 3).

The differences in results with lagged participation included and excluded may be interpreted as indicating that the impacts of the explanatory variables on the decision to participate of women in rural areas are very different from those in urban areas. However, once a woman decides to participate or not, the decision to continue to participate or to remain out of the workforce is the same in both rural and urban areas.

Reduced Form Wage and Hours Equation

In order to test for sample selection bias in the structural regime, the inverse mills ratio (λ_i) was calculated and entered into reduced-form regressions of the wage and hour equations. These regressions included slope shifters for non-SMSA residence. The results, presented in Table 4, show that selection bias does not exist in either the hours equation or the log wage regression.[7] Because selection bias is not deemed to be a problem, the (log) wage and hours reduced forms were then estimated by OLS (i.e., without a selectivity correction) and checked for the statistical assumptions of normality and homoskedasticity (Spanos). Normality was

tested using a Bera-Jarque Skewness-Kurtosis test described on pp. 453-454 of Spanos. Heteroskedasticity was tested using White's test described in pp. 466-467 of Spanos. The OLS reduced form wage and hours equations passed tests of normality, and showed no signs of heteroskedasticity.[8]

Structural Wage Equation

The structural wage equation (equation 7)was estimated with predicted hours used as an instrument to identify the endogenous hours worked. The initial imposed identifying restriction is $\underline{\alpha}_4^u = \underline{0}$ (see equation 7), or that Z_i^* only affects labor demand through its impact on hours worked (supply). Next, Basmann tests of over-identifying restrictions $\{\underline{\alpha}_1^r = \underline{0}, \underline{\alpha}_2^r = \underline{0}, \underline{\alpha}_4^r = \underline{0}\}$ were conducted, to see if rural-urban effects on the demand side could be fully captured by an intercept shifter.[9] This hypothesis that rural-urban effects on the demand side can be fully captured by an intercept shifter could not be rejected $\{F(21, 2172) = .6150\}$. Thus, the effects of rural-urban differences in labor demand were captured in two different ways. First, any systematic rural-urban difference in labor supply by married females was captured during the identification of the endogenous hours variable, whereas rural-urban differences in labor demand were captured by an intercept shifter and by any variation in the response to hours from rural to urban areas. The t-test for the coefficient on $h_i * R_i$ indicated that there is no differential demand response to hours from rural to urban areas. This variable was then deleted.

The final 2SLS estimation of the wage equation is presented in Table 5. Model 1 contains the estimates with $h_i * R_i$ included, and Model 2 contains the version with this variable excluded. Once the demand response to hours worked is held constant from rural to urban areas, the impact of SMSA vs. non-SMSA status (R_i) becomes significant. The model results show clearly that labor demand differences between urban and rural areas can be captured by the intercept shifter (non-SMSA) which was highly significant.[10]

---

[7]The results in Table 4 were calculated using the "Heckit" 2-step procedure (see Heckman 1979 and Maddala, pp. 231-236). The reported standard errors are corrected as Heckman recommended. This correction procedure is described in Heckman (1979) and Madalla. The significance of the λ_i in Table 4 determines whether selectivity bias is important. Beyond this, λ_i has no real significance in this model.

[8]These results are available from the authors on request.

[9]The basic logic behind the Basmann tests follows. The model is estimated using 2SLS with all the over-identifying restrictions imposed. The residuals from this regression (calculated using the *actual* endogenous right-hand side variable) are then regressed on the entire instrument set (with the predicted log (wage) and predicted hours not included). Tests are then conducted for the exclusion of the over-identifying variables from this second regression. See Mroz for a similar application of the Basmann test.

[10]Alternatively, R_i could have been deleted and the $h_i * R_i$ included. Both sets results (with R_i in and $h_i * R_i$ out; and with R_i out and $h_i * R_i$ included) indicate strong differences in demand from rural to urban areas.

Table 4. Reduced form wage and hours equation used to test for sample selection bias[a]

Variable	Dependent Variable ln Wage	Hours
Intercept	.568 (2.37)	1,507.55 (4.42)
Child < 1	.063 (1.02)	-85.9 (-.98)
Child 2-6	.051 (1.98)	-189.64 (-5.20)
Child 7-12	-.036 (-2.04)	-77.08 (-3.04)
Child 13-18	-.026 (-1.57)	-15.26 (-.66)
Husband's Earnings[b]	.046 (3.77)	-62.74 (-3.64)
Unearned Income[b]	.072 (2.50)	-77.53 (-1.89)
Age	.025 (3.76)	17.30 (1.83)
Age Squared	-.00025 (-3.01)	-.28 (-2.42)
Highest Grade	-.031 (-1.09)	43.04 (1.08)
Highest Grade Squared	.306 (3.07)	-.86 (-.61)
Race	-.014 (-.44)	-18.69 (-.43)
Husband's Highest Grade	.010 (2.11)	-5.93 (-.89)
Non-SMSA	.412 (1.07)	394.77 (.72)
λ	-.985 (-.88)	-113.29 (-.72)
State Effects[c]		
Maryland	.120 (2.80)	-59.23 (-.97)
Virginia	.039 (.91)	69.01 (1.12)
North Carolina	.-.085 (-1.78)	115.24 (1.69)
Georgia	.-.019 (.39)	-18.45 (-.27)
Florida	-.054 (-1.32)	72.58 (1.26)
Kentucky	-.047 (-.83)	-109.95 (-1.37)
Tennessee	-.020 (-.40)	-.87 (-.01)
Alabama	-.021 (-.40)	22.24 (.30)
N	2,216	2,216
$\bar{R}^2$	.221	.059

[a]Non-SMSA-variable interactions not shown to conserve space. T-statistics (in parentheses) are computed using corrected standard errors (see footnote 7).
[b]In $10,000
[c]Deleted state fixed effect is Louisiana.

Table 5. Two-Stage Least Squares Estimates of Wage Equation[a]

	Dependent Variable = log (wage)	
Explanatory Variable	Parameter Estimates Model 1	Parameter Estimates Model 2
Intercept	1.16 (4.96)	1.16 (5.24)
Hours[b]	-.0277 (-3.23)	-.0279 (-3.59)
R * Hours[b]	-.00097 (-.07)	- -
Age	.0266 (5.26)	.0265 (5.30)
Age Squared	-.00029 (-4.58)	-.00029 (-4.62)
Highest Grade	-.03593 (-1.56)	-.03591 (-1.56)
Highest Grade Squared	.00351 (4.24)	.00351 (4.24)
Race	-.0433 (-1.75)	-.0432 (-1.75)
Husband's Highest Grade	.00834 (2.17)	.00835 (2.18)
Non-SMSA	-.05791 (-.22)	-.0761 (-3.93)
State Effects[d]		
Maryland	.09501 (2.30)	.0949 (2.30)
Virginia	.03529 (.90)	.03532 (.90)
North Carolina	-.04383 (-1.14)	-.0436 (1.14)
Georgia	-.0252 (-.65)	-.02534 (-.65)
Florida	-.05344 (-1.46)	-.05301 (-1.47)
Kentucky	-.05831 (-1.30)	-.0578 (-1.31)
Tennessee	-.04664 (-1.12)	-.0464 (-1.12)
Alabama	-.05585 (-1.28)	-.05591 (-1.28)
N	2,216	2,216
$\bar{R}^2$	.077	.078

[a]T-statistics in parentheses.
[b]Endogenous variable in 100 hours annually.
[c]Tests for structural differences from rural to urban areas failed to reject over-identifying restrictions.
($\alpha^r_1 = 0$, $\alpha^r_2 = 0$ $\alpha^r_4 = 0$), $F(20,2175) = .617$.
[d]Louisiana is the deleted state effect.

Demand differences between states are captured in both models by the fixed state effects, which are jointly significant at the .05 level. The relatively low R^2 is not unusual with cross-sectional data (see, for example, Tokle and Huffman). The variables in Z_i all have their expected signs. Wages increase with age up to age 46 when they begin to decline; this concave profile is consistent with other studies (e.g., Huffman and Lange, who use experience rather than age). Education has a positive effect; though the

linear term is negative, it is not significantly different from zero, and the squared term has a significant positive effect. Employers pay significantly more for more highly educated employees. Using the point estimates of the coefficients, additional education beyond grade five leads to higher wages. Race does not have a significant impact on wages. The highly significant rural/urban dummy shows that at the mean wage, rural married female workers earn 7.6 percent lower hourly wages than urban workers.

The negative sign on the hours variable (h_i) in this equation is an interesting result. Some argue that employers will be willing to provide a premium wage for full-time workers, and thus, the sign of this variable should be expected to be positive. Blank (1990) found using CPS data and controlling for the choice of full versus part-time employment, that part-time workers received higher wages than full-time workers. The evidence here is that full-time wages are, on average, lower. There are several plausible explanations. Full-time employees receive benefits other than wages, and thus, their money wages might be lower. Second, once the determinants of hours supplied and labor force participation along with the human capital determinants of wages are controlled for, firms pay less for full-time workers whose marginal productivity might be lower. A final explanation is that most of the variation in hours worked is among workers who already might be classified as part-time workers, and among part-time workers, premiums are paid for people who are willing to work short hours on flexible schedules. These issues are beyond the scope of this study, but are certainly subjects for future studies. The negative coefficient on hours suggests that in the southern United States, women are willing to accept lower wages for longer hours worked.

The estimation of the wage equation provides insight about how wages vary between rural and urban areas. This variation can be fully captured by differences, if any, in labor supplies and by a single intercept shifter. There is no independent variation in labor demand across rural/urban designation between states (tested by α_4''), nor are there differences in slopes of the other variables (Z_i, Z_i^*).

Structural Supply Equation

The first specification used for the hours supplied was: hours = f (Z^*, log (wage) log (wage) * R_i, R_i). The log of the wage and its interaction with the SMSA/non-SMSA dummy are endogenous variables in the hours equations (see the simultaneous system (3) and (4)); therefore, an identifying restriction was needed. The exogenous variables excluded from this estimation was Z_i and the interactions between R_i and Z_i^* and S_i. S_i was also excluded from this first regression (i.e., supply was not allowed to vary by state of residence). The exclusion of S_i from the supply equations was the first overidentifying restriction tested. The test statistic (F(8,2175) = 2.44, $F_{c,.05}$ (8,2175) = 1.94) indicated strong rejection. Thus, S_i was entered as a supply determinant.

The results from this model, where supply was allowed to vary by state of residence, (Model 1) are shown in the first column of Table 6. A Basmann test of the over-identifying restrictions ($b_2r = \underline{0}$, $b_3r = \underline{0}$, $b_4r = \underline{0}$) used in Model 1, leads to a marginal rejection (F(20,2175) = 1.62, critical value = 1.57) of these restrictions.[11] This rejection provides some evidence of a systematic difference in labor supply behavior between rural and urban areas.

In order to further examine the source of these systematic supply differences, Basmann tests of individual groups of coefficients were conducted. A Basmann test of the over-identifying restriction that the coefficients on Z_i^* do not vary from rural to urban areas (i.e., $\underline{b_2}r = 0$) lead to rejection of the restriction (F(6,2175) = 2.53, critical value = 2.10). The model was then respecified with the Z_i^* coefficients allowed to vary across the rural-urban designation. These results are presented as Model 2 in Table 6.

Further tests of over-identifying restrictions ($b_3r = \underline{0}$; $\underline{b_4}r = \underline{0}$) failed to reject the null hypothesis of no differences between urban and rural areas for the variables Z_i and S_i respectively. Because no additional over-identifying restrictions could be rejected, Model 2 was selected as the best model.[12] A test for deletion of R_i * ln(wage) failed to reject; the column labeled Model 3 shows the estimate with this variable deleted. The evidence is that there are significant differences in female labor supply behavior across rural and urban areas, and that these differences cannot be explained solely by differences in wages. These differences also cannot be captured by a simple intercept shifter; the slopes of the coefficients of Z_i^* vary across the rural/urban designation.

The parameter estimates resulting from Model 3 are broadly consistent with those from other studies, except that the supply determinants (Z_i^*) show variation from rural to urban areas. The existence of children in the family leads to fewer hours supplied,

---

[11] This test examines whether $R_i * Z_i^*$, $R_i * S_i$, and $R_i * Z_i$ can be used as instruments for log (wage). If they cannot, then they, or some subset of them should be used as regressors in the hours supplied equation.

[12] The result that $\underline{b_2}r \neq \underline{0}$ was robust to the order of testing of the three restrictions ($\underline{b_2}r = \underline{0}$, $b_3r = \underline{0}$, $\underline{b_4}r = \underline{0}$).

Table 6. Two-stage Least-squares Results for Hours Supplied[a]

Explanatory Variable	Model 1[b]	Model 2[c]	Model 3[d]
Intercept	1,757.98	1,749.02	1,727.71
	(13.40)	(13.06)	(14.74)
1n(wage)[e]	173.94	206.33	222.64
	(1.94)	(2.20)	(2.80)
R_l * 1n(wage)[e]	168.90	55.73	-
	(1.21)	(.33)	-
Child < 1	-56.98	-43.87	-44.82
	(-.94)	(-.54)	(-.56)
Child 2-6	-130.48	-157.41	-157.83
	(-4.89)	(-4.48)	(-4.49)
Child 7-12	-1.16	-40.16	-39.58
	(-.06)	(-1.62)	(-1.60)
Child 13-18	16.66	1.20	1.56
	(.06)	(.06)	(.07)
Husband's Earnings[f]	-54.69	-74.66	-76.48
	(-3.58)	(-4.06)	(-4.36)
Unearned Income[f]	209.51	-177.84	-179.57
	(-6.08)	(-4.53)	(-4.62)
Non-SMSA	-230.05	-186.07	-117.64
	(-1.18)	(-.87)	(-2.22)
State Effects			
Maryland	-104.94	16.95	-102.15
	(-1.87)	(.27)	(-1.84)
Virginia	55.97	.80	54.39
	(1.08)	(.01)	(1.05)
North Carolina	58.10	75.40	61.86
	(1.13)	(1.09)	(1.22)
Georgia	29.43	-46.35	25.74
	(.56)	(-.66)	(.49)
Florida	9.45	69.77	17.59
	(.20)	(1.19)	(.36)
Kentucky	-114.88	2.85	-106.19
	(-1.96)	(.04)	(-1.82)
Tennessee	-8.29	14.93	-8.58
	(-.15)	(.21)	(-.16)
Alabama	16.43	-11.79	13.94
	(.28)	(-.164)	(.24)
R * child < 1	—	-27.12	-24.58
		(-.22)	(-.20)
R * child 2-6	—	69.75	72.18
		(1.30)	(1.36)
R * child 7-12	—	91.91	90.42
		(2.43)	(2.41)
R * child 13-18	—	-3.22	-4.02
		(-.10)	(-.12)
R * husband's earnings	—	63.47	69.45
		(1.92)	(2.51)
R * unearned income	—	-101.76	-94.60
		(-1.26)	(-1.21)
N	2,216	2,216	2,216
$\bar{R}^2$	.048	.062	.066

[a] T-Stastistics are in parentheses. Louisiana is the deleted state effect.
[b] Over-identifying restrictions imposed for Model 1 are coefficients on Z, Z*, and state residence, do not vary from rural to urban areas. Test of these restrictions $F(20,2175) = 1.52$. Thus, the restriction are marginally rejected.
[c] Over-identifying restrictions imposed for Model 2 are coefficients on Z, and state fixed effects do not vary from rural to urban areas. A test of over-identifying restrictions involving Z* leads to $F(6,2175) = 2.53$. Thus, the restrictions is rejected.
[d] Over-identifying restrictions imposed for Model 3 are that the coefficients on Z and state fixed effects do not vary from rural to urban areas. None of these restrictions were rejected.
[e] Endogenous variable.
[f] In $10,000.

income effects are strongly negative, indicating that leisure is a normal good, and the wage effect falls somewhere in the middle of the studies cited in Mroz. The interesting part of the results relates to the differences in labor supply parameters from rural to urban areas. These differences can help to identify constraints to labor supply, or differences which rural areas might face. First, women in non-SMSA areas supply significantly fewer hours than those from SMSA.

A second difference emerges from the positive sign of the coefficient associated with the $R_i * Child\ 7\text{-}12$ variable. The positive sign of this coefficient shows that the presence of children of these ages has a smaller negative effect on hours supplied in rural areas than in urban areas. In fact, the magnitude of the parameter estimate for $R * Child\ 7\text{-}12$ shows that children of this age group actually have a positive effect on married female labor supply in rural areas. There are a number of plausible explanations for this, one being that child care is less of a constraint to labor supplies in rural areas than in urban areas. This needs to be examined more carefully in a study that contains information on the distance to child care facilities; these data are not available in the CPS.

The differences between income effects across rural and urban areas also merit discussion. Husband's earned income has virtually no effect on labor supply in rural areas, while it has a strongly negative effect in urban areas. The net effect of total family income (husband's earned plus family unearned income) on female labor supply in rural areas is dampened to a large degree by this difference. Given the lower average household income in rural areas, this weaker income effect is consistent with the observation that lower income households are less likely to lower their labor supplies as income rises than are higher income households. A second explanation is that there are fewer activities to substitute for women's work time in rural areas. Without these alternatives, wealthier women continue to work in rural areas. Both of these can be investigated in future research.

SUMMARY AND CONCLUSIONS

This paper provides evidence on the labor supply of married females in the southern U.S. There are differences in this behavior between rural and urban areas. Some of these differences are caused by demand side effects, but differences in supply behavior also exist. There was no measurable difference in differential participation rates between urban and rural areas though the parameters affecting overall participation were different between these areas. Demand differences were found to be captured by an intercept shifter, and supply differences included some slopes that varied from urban to rural areas.

The differences in labor supplies were manifested in several ways. First, the existence of children between 7 and 12 years old had a positive effect on labor supply in rural areas. In urban areas these children caused fewer hours to be supplied. Child care may be less of a constraint in rural areas for several reasons. Manufacturing, a major employer in rural areas, may allow couples to work alternate shifts. Because rural families are less mobile, other family members may be nearby to provide child care. Additionally, given lower rural crime rates, rural families may feel more secure about latch-key children. Finally, when their children are school age, rural women may choose to work in order to decrease isolation.

Second, income effects were weaker in rural areas. Husband's earnings are a strong substitute for wives' earnings in urban areas. This is not true in rural areas. The absence of enriching alternative uses of women's time in rural areas may help explain this, though more research along these lines is needed. Family unearned income has a more negative impact on women's hours in rural areas than in urban areas (Table 6). Even though this difference is not statistically significant, it, combined with differences in responses to earned income, shows differences in how rural and urban families react to unearned and earned income.

Third, the state of residence has strong effects on labor force participation by women. These effects may include variation in labor laws, differing industrial mixes and differential access to job training and child care. Though these variations are not examined in detail in the study, the significant fixed state effects show some major differences. Labor demand and labor supply also differed across the states in the South. Particularly, labor supply varied significantly from state to state; state policies can affect female labor supplies.

The study outlines several new areas of research. The data do not permit use of location-specific variables that show variation within the state. County-level unemployment rates, child-care facilities, educational opportunities, etc., all may affect participation, demand, and supply. The impacts of these variables should be examined. The effects of fixed costs of work, job search, etc. also could not be examined given data limitations. These variables all might have a different effect on labor markets in rural versus urban areas. Their effects should be examined within a framework similar to that used in this study.

The paper shows that studies of labor supplies need to consider the obvious potential differences between labor supplies of married females in rural areas and urban areas. Careful consideration should be given to these differences, and the method of empirical analysis should be flexible with respect to these differences.

REFERENCES

Basmann, R. L. "On Finite Sample Distributions of Generalized Classical Linear Identifiable Test Statistic." *J. Am. Stat. Assoc.*, 55(1960): 650-659.

Blank, Rebecca M. "Are Part-time Jobs Bad Jobs?," In *A Future of Lousy Jobs?* Gary Burtless, ed. Washington, DC: Brookings Institution, 1990, pp. 123-164.

Blank, Rebecca M. "Simultaneously Modeling the Supply of Weeks and Hours of Work Among Female Household Heads." *J. Labor Econ.*, 6.2 (1988): 177-204.

Blau, David M., and Philip K. Robins. "Child-care Costs and Family Labor Supply." *Rev. Econ. and Stat.*, (1988): 374-381.

Blundell, Richard, John Ham, and Costas Meghir. "Unemployment and Female Labor Supply." *The Economic Journal*, 97(1987): 44-64.

Heckman, James, "Shadow Prices, Market Wages, and Labor Supply." *Econometrica*, 42(1974): 679-894.

Heckman, James. "The Common Structure of Statistical Models of Truncation, Sample Selection, and Limited Dependent Variables and a Simple Estimator for Such Models." *Annals of Econ. and Soc. Meas.*, 5(1976): 475-492.

Heckman, James. "Sample Selection Bias as a Specification Error." *Econometrica*, 47(1979): 153-161.

Huffman, Wallace E., and Mark D. Lange. "Off-farm Work Decisions of Husbands and Wives: Joint Decision Making." *Rev. Econ. and Stat,.* (1989): 471-480.

Lundberg, Shelly. "Labor Supply of Husbands and Wives: A Simultaneous Equations Approach." *Rev. Econ. and Stat.*, (1988): 224-235.

Maddala, G. S. *Limited-dependent and Qualitative Variables in Econometrics.* New York: Cambridge University Press, 1983.

Mroz, Thomas A. "The Sensitivity of an Empirical Model of Married Women's Hours of Work to Economic and Statistical Assumptions." *Econometrica.* (1987):765-799.

Nakamura, Alice, and Masao Nakamura. "A Comparison of the Labor Force Behavior of Married Women in the United States and Canada with Special Attention to the Impact of Income Taxes." *Econometrica*, 49.2 (1981): 451-489.

Nakamura, Alice, and Masao Nakamura. "Dynamic Models of the Labor Force Behavior of Married Women which can be Estimated Using Limited Amounts of Past Information." *J. Econometrics.* 27(1985): 273-298.

Nakamura, Masao, Alice Nakamura, and Dallas Collen. "Job Opportunities, the Offered Wage, and the Labor Supply of Married Women." *Am. Econ. Rev.*, (1979): 787-805.

Spanos, Aris. *Statistical Foundations of Econometric Modelling.* New York: Cambridge University Press, 1986.

Stern, Nicholas. "On the Specification of Labor Supply Functions." *Unemployment, Search and Labor Supply,* In Richard Blundell and Ian Walker, eds., New York: Cambridge University Press, 1986. pp. 143-190.

Summers, Gene F., Francine Horton, and Christian Gringeri. "Rural Labor Market Changes in the U.S." In *National Rural Studies Committee: A Proceedings*, Emery N. Castle, ed. Cedar Falls, Iowa, May 1990, pp. 61-79.

Tokle, J. G., and Wallace E. Huffman. "Local Economic Conditions and Wage Labor Decisions of Farm and Rural Nonfarm Couples," *Am. J. Agr. Econ.*, 73 .3(1991): 652-670.