Appendix A: Modeling Rural County Socioeconomic Change

We measure population and job growth as the natural log of the ratio of the number of people (jobs) in each county in 1992 or 2000 relative to 1985.[70] In modeling rural growth, a county's historic pattern of population and employment change are often key predictors. County changes in population and employment are included for both the 1970s and the years immediately preceding the introduction of the CRP (1982-85). In the 1970s, agriculture, mining, and manufacturing were all relatively prosperous and contributed to the rural rebound of the period. In contrast, these industries suffered in the 1980s. The inclusion of 1982-85 changes captures some of this decline. As with the dependent variable, these explanatory variables take the log form.

A series of demographic variables captures the effects of race, ethnicity, age, educational attainment, and population density on the community growth process. All were from the 1980 Census of Population, with all but population density expressed as a percentage of the total. Population density entered the equations in log form. To measure scenic attractiveness, the presence of high mountains (0/1 dummy variable), the prominence of surface water (in log form) and forests (percentage of land area) are included in analyses of the entire study group. Also included are z-scores of several climate measures (McGranahan, 1999). For the matched-pair analysis, these amenity measures were replaced by the "natural amenity scale" developed by McGranahan to combine all of these factors into one measure. Table A.1 provides the mean values of the employment, demographic, and amenity variables considered.

Measures of initial industry structure are ubiquitous in studies of job growth. Industry structure is measured by the proportion of employed residents working in agriculture, manufacturing, mining, business services (finance, insurance, real estate, and other professional services), and recreation (eating places, amusement, and recreation, other than hotels) in 1980. Somewhat unique rural industrial expansion in the 1980s was from casinos, prisons, and large meatpacking plants. To take account of the sometimes dramatic changes accompanying these developments, dummy variables were included to reflect whether a county had any of these industries in 2000.

Local labor market and locational characteristics could also affect community growth. Higher employment rates and higher incomes (in log form) might encourage migration, but might discourage new employers. The attractiveness of an area is a function of its access to services and other amenities, measured by whether the county was adjacent to a metropolitan area in 1983 (represented as a 0/1 dummy variable). The growth potential of a county may also depend on the percentage of its residents commuting outside the county to work. Finally, because the Great Plains has its own unique characteristics, a dummy variable indicates whether or not the county was in the Great Plains.

[70] We considered modeling net migration, but intercensal net migration estimates are not available. Furthermore, the small populations of counties studied make the reliability of any intercensal population estimates suspect. Independent measures of elderly and children were included in the population change analysis to reflect their influence on population trends due to age structure and fertility.

Table A.1—Mean values of demographic and employment trends and amenity variables

Variable description	Unit	Study counties	High-CRP[1]	Matched counties
Dependent variables				
Post-CRP population change:				
1985-1992 (short run)	Percent	-1.5	-9.0	-5.9**
1985-2000 (long run)	Percent	5.3	-9.8	-4.1**
Post-CRP employment change:				
1985-1992 (short run)	Percent	5.6	-3.7	1.4**
1985-2000 (long run)	Percent	23.9	7.6	13.4**
Explanatory variables				
Pre-CRP population and employment change:				
1970-1982 population	Percent	11.3	-3.2	3.3**
1982-1985 population[2]	Percent	-0.3	-2.3	-1.3**
1970-1982 employment	Percent	17.6	1.6	13.5**
1982-1985 employment[2]	Percent	2.6	-1.7	0.3**
Demographic characteristics:				
Black population, 1980	Percent	7.1	0.6	0.4
Hispanic population, 1980	Percent	4.2	4.4	6.9
Native American population, 1980	Percent	1.5	3.3	1.9
Population under 18, 1980	Percent	29.7	29.8	29.3
Population over 62, 1980	Percent	18.2	19.3	19.7
Under 12 years of school, aged 25-44, 1980	Percent	23.4	17.2	16.5
College grads, aged 25-44, 1980	Percent	14.1	16.9	17.4
Population density, 1980	P/sq mi	24	5	10**
Natural amenity characteristics:				
High mountains dummy variable	0/1[3]	7.4	5.6	10.8
Water/total area (x 10)	Log	-2.1	-6.5	-6.2
Land in forest	Percent	26.7	3.7	8.5**
January days with sun (x 10)	Z-score[4]	1.8	5.2	5.4
January temperature (x 10)	Z-score	-1.9	-8.3	-6.1*
July humidity (x 10)	Z-score	2.3	9.7	7.1**
July temperature (x 10)	Z-score	-2.6	-4.8	-5.0
Natural amenities scale (x 10)	Z-score	-3.6	-7.2	-6.6

[1]High-CRP counties have an average CRP rental-payment-to-income ratio for 1991-93 exceeding 2.75 percent. Of the 1,481 study counties, 195 were high-CRP by this definition.
[2]We include 1982-85 trends separately because rural county growth was slower in this period than during the preceding 12 years.
[3]Set to one if mountains are present. The data represent the proportion of observations coded "1."
[4]Z-scores are the number of standard deviations an observation differs from the mean (across all observations).
Source: BEA Income files, 1980 Census of Population and McGranahan (1999).
* and ** indicate that the difference between high-CRP counties and their matched pairs is significantly greater than 0 at the 0.05 and 0.01 level, respectively.

Because CRP primarily affects farming-dependent areas, several agricultural variables in addition to employment were included in the analysis. Most of these are from the 1982 Census of Agriculture. Finally, the ratio of CRP enrollment to total cropland or the ratio of CRP rental payments to county household income is included to measure CRP's local importance. Mean values of the industry and farm structure variables are presented in table A.2.

Our database includes over 45 measures that have been associated with population and job change or that reflect local agricultural conditions. While these explanatory variables should capture the independent effects of many county characteristics potentially related to population or employment

Table A.2—Mean values of industrial, labor market, and farm structure variables

Variable description	Unit	Study counties	High-CRP[1]	Matched counties
Local economic characteristics:				
Agricultural employment, 1980	Percent	16.7	31.7	24.7**
Manufacturing employment, 1980	Percent	17.6	5.7	8.4**
Mining employment, 1980	Percent	2.5	2.2	2.3
Business services employment, 1980	Percent	4.2	3.9	4.2*
Recreation employment, 1980	Percent	4.1	4.1	4.5*
Special development dummy variables[2]:				
Prison county	0/1	2.6	1.0	0.0
Casino county	0/1	0.9	0.0	1.5
Meatpacking plant county	0/1	1.4	0.5	1.0
Labor market and location characteristics:				
Civilian employment, age 15-64, 1980	Percent	62.7	64.9	65.6
Working outside the county, 1980	Percent	19.0	10.9	12.9*
Median household income, 1979	$	12,840	12,620	12,936
Adjacent to a metropolitan area, 1983	0/1[2]	41.3	15.9	22.6
Great Plains county	0/1[2]	27.1	80.0	59.5**
Agricultural characteristics:				
Cropland/all land, 1982	Percent	40.5	46.7	45.1
Irrigated farmland, 1982	Percent	4.5	4.3	8.5**
Grain/total sales value, 1982	Percent	29.5	38.4	31.5**
Wheat/total sales, 1982	Percent	8.8	25.2	12.2**
Livestock/total sales, 1982	Percent	56.2	51.5	61.6**
Govt. payments/total income, 1981-83	Percent	1.6	6.0	2.6**
CRP enrollment/cropland, 1991-93	Percent	8.0	21.3	5.1**
CRP payments/income, 1991-93	Percent	1.3	6.7	0.8**
Farm sales/household income, 1980	Percent	0.8	1.9	1.4**
Farms w/ sales over $250,000 in 1982	Percent	4.7	5.3	5.8
Farms w/ sales under $20,000 in 1982	Percent	51.5	35.7	38.9*
Farmers working off-farm 200+ days, 1982	Percent	28.0	17.9	21.0**

* and ** indicate that the difference between high-CRP counties and their matched pairs is significantly greater than 0 at the 0.05 and 0.01 level, respectively.

1 High-CRP counties have an average CRP rental-payment-to-income ratio for 1991-93 exceeding 2.75 percent. Of the 1,481 study counties, 195 were high-CRP by this definition.

2 The data reported for all 0/1 dummy variables represent the percentage of observations coded "1" rather than the mean for expositional ease.

Source: 1980 Census of Population, 1982 Census of Agriculture, BEA Income file, and CRP Contracts file.

change, several socioeconomic measures are highly correlated, with no *a priori* reason for selecting one over the other. To avoid statistical problems from estimating relationships with an over-identified model, in addition to a standard analysis including all explanatory variables, a backward stepwise regression procedure is used to narrow the set of explanatory variables.[71]

For the matched-pair analysis, two versions of the model were estimated using a subset of explanatory variables. In the first, CRP measures were excluded from the equation, leaving the constant term to capture CRP's impact on the difference in growth trends between high- and low-CRP counties. We estimated a second set of regressions with the difference in the CRP measure between matched pairs included as an independent variable, with the constant constrained to equal zero, to capture the impact that varying levels of CRP participation had on socioeconomic trends.

[71] In this procedure, the socioeconomic measure with the least statistical significance is removed and the analysis is repeated until all remaining measures are statistically significant. In using this procedure, we exempt the CRP measures from possible elimination and exclude other measures not significant at the 10-percent level. This approach biases the analysis in favor of finding a significant relationship between CRP use and socioeconomic trends.

In addition to including mining employment in 1980 as an explanatory variable, we also created a separate set of matched pairs that excluded counties where mining comprised over 5 percent of 1980 employment. Doing so helped clarify the relationship between community development and CRP, since variations in mining added a lot of statistical "noise" to the data.

Much of the employment-migration literature recognizes that population and employment growth rates are endogeous phenomena to be modeled simultaneously. The focus of this literature has been on whether employment growth is stimulated by the in-migration of people drawn to an area by quality of life considerations. In our analysis, we are not concerned with the mechanisms through which the CRP program might have affected population or employment; we are concerned with overall effects. Recognizing the inherent simultaneity, we use the same independent measures in both the employment and population equations. Our analyses are therefore equivalent to a reduced form equation from a simultaneous equation model.

The benefits of the matched-pair approach are its intuitive appeal, transparency, and the fact that it is less dependent on assumptions regarding functional forms of structural models or even reduced-form relationships.[72] That is, because the matched pairs are relatively "close," there is less need for controls; and the use of a linear model to control for potentially convoluting factors should give a reasonable approximation of even nonlinear effects, because the differences in explanatory variables are relatively small.

The quasi-experimental control group approach we adopt builds on analysis of experimental data in that it attempts to assess the impact of a "treatment" by developing an appropriate counterfactual. When the treatment is randomly distributed within the population being studied, the "control" group is implicitly all observations that haven't been treated. But when the treatment is not randomly distributed, selection of a control group indicating how treated observations would have developed in the absence of treatment becomes a little more difficult. In such cases, the development of all nontreated observations may not be the appropriate counterfactual.

Developing an appropriate control group is at the heart of quasi-experimental control group analysis. There are many ways to operationalize the control group concept – matches can be one-to-one, one-to-many, or many-to-many; they can be based on nearest neighbor, by an *ad hoc* comparison of one or two key characteristics, or by using a statistical measure of similarity, such as a propensity score or the Mahalanobis distance. We have adopted the matched-pair (one-to-one) technique based on minimizing the overall Mahalanobis distance used by Isserman and Rephann (1995) because of its flexibility and its intuitive appeal. By applying a difference-in-differences analysis to observations that have been matched on the basis of growth factors, the approach adopted here should highlight CRP's potential impacts on economic trends (Blundell and Diaz, 2000).

The most complicated growth model estimated for this report examines the interaction between population density and CRP enrollment. The model attempts to explain differences in job growth trends between high-CRP counties and their matches as a function of differences in a series of explanatory variables, based on counties where mining accounted for 5

[72] By comparison, recovering the structural components of a simultaneous equations model is much more difficult and requires much stronger assumptions. To do so, one needs to be able to justify both the functional form and at least two exclusionary restrictions: what exogenous variables influence employment growth but do not influence population growth (or net migration), and vice versa. These "instruments" would also need to be uncorrelated with unobservable variables affecting the other equation.

percent or less of employment in 1980. This analysis was used to construct the CRP impacts presented in figure 3.3 in the text. Table A.3 presents the regression results for the full model explaining job growth (neither CRP nor its interaction with population density were significant in the population growth model, so the results are not reported). The backward stepwise regression results were very similar, although a couple additional control variables were identified as having a statistically significant impact on job growth (all the significant variables from both the full and backward stepwise regressions are in bold in table A.3).

Finally, farm-related enterprises were identified to explore the extent to which this group of businesses was particularly susceptible to changes in CRP enrollment. Table A.4 lists the 3- and 4-digit SIC codes for industries

Table A.3—Interaction between population density and CRP's impact on job growth

Explanatory variables (low-CRP minus high-CRP value)	Unit	Short-term job growth (1985-1992)		Long-term job growth (1985-2000)	
		Beta	t-statistic	Beta	t-statistic
CRP payments to income ratio	Percent	0.085	0.735	**0.236**	**1.923[a]**
Population density, 1980	log	0.011	0.069	0.195	1.134
Density x CRP ratio	Percent	**-0.216**	**-2.169***	-0.075	-0.715
Employed in ag, 1980	Percent	**-0.455**	**-3.369****	-0.161	-1.125
Density x Percent ag emp.	Percent	0.079	0.683	-0.030	-0.243
Population, 1982/1970	log	-0.081	-0.716	0.161	1.340
Population, 1985/1982	log	0.108	1.456	0.062	0.794
Employment, 1982/1970	log	-0.055	-0.556	**-0.216**	**-2.069***
Employment, 1985/1982	log	-0.071	-0.981	-0.076	-0.981
Under 18 years of age, 1980	Percent	0.193	1.378	0.178	1.195
Over 62 years of age, 1980	Percent	-0.098	-0.712	0.010	0.069
American Indian, 1980	Percent	-0.002	-0.020	0.104	0.999
Black, 1980	Percent	**-0.181**	**-2.676****	**-0.230**	**-3.212****
Hispanic, 1980	Percent	-0.044	-0.419	0.091	0.810
Cropland, 1982	Percent	-0.180	-1.545	-0.156	-1.262
Livestock/total sales, 1982	Percent	-0.031	-0.450	0.023	0.314
Govt payments/income, 1981-83	Percent	0.005	0.039	-0.122	-0.981
Wheat/total sales, 1982	Percent	-0.134	-1.530	-0.069	-0.744
Less than high school, 1980	Percent	-0.006	-0.051	-0.130	-1.010
College, 1980	Percent	0.119	1.520	0.044	0.526
Civilian employment rate, 1980	Percent	0.004	0.046	0.059	0.663
Median household income, 1979	Dollars	**-0.198**	**-1.899[a]**	-0.079	-0.712
Natural amenities index	Z-score[1]	0.036	0.462	-0.039	-0.464
Land in forest	Percent	0.066	0.664	**0.261**	**2.482***
Great Plains county	0/1[2]	**-0.156**	**-1.885[a]**	-0.139	-1.585
Employed in mining, 1980	Percent	**-0.199**	**-2.999****	-0.072	-1.021
Employed in recreation, 1980	Percent	0.019	0.256	0.031	0.394
Commuting outside county, 1980	Percent	0.018	0.239	0.062	0.773
Meat packing plant county	0/1	0.052	0.843	0.026	0.398
Casino county	0/1	0.027	0.470	0.069	1.117
Prison county	0/1	-0.052	-0.832	-0.022	-0.332
Adjusted R-squared		0.41		0.34	

[a], *, and ** indicate the regression coefficient is statistically different from 0 at the .10, .05, and .01 level of significance, respectively. Bold indicates variables that were significant at the .10 level or lower in this or the backward stepwise regressions. Beta represents the standardized regression coefficient. Adjusted R-squared indicates the portion of variation explained by the regression.
[1]Z-scores are the number of standard deviations an observation differs from the mean (across all observations).
[2] Dummy variables with a "0" or a "1" value.

Source: Economic Research Service calculations using data from the 1980 Census of Population, the 1982 Census of Agriculture, the Bureau of Economic Analysis, and the CRP Contracts file. Matched pairs exclude counties with more than 5 percent employed in mining. The constant term was constrained to equal 0.

Table A.4—Agricultural services industries

SIC code[1]	Description
071	Agricultural services: Soil preparation services
072	Agricultural services: Crop services
076	Agricultural services: Farm labor and management services
1542	Construction: Nonresidential construction, NEC
203	Food products: Canned, frozen, and preserved fruits, vegetables, etc.
204	Food products: Grain mill products
2061	Food products: Cane sugar, except refining
2062	Food products: Cane sugar refining
2063	Food products: Beet sugar
2074	Food products: Cottonseed oil mills
2075	Food products: Soybean oil mills
2076	Food products: Vegetable oil mills
4212	Transportation: Local trucking, without storage
4221	Transportation: Farm product warehousing and storage
4449	Transportation: Water transportation of freight, NEC
4731	Transportation services: Freight and cargo
5083	Wholesaling: Farm and garden machinery and equipment
5153	Wholesaling: Grain and field beans
5159	Wholesaling: Farm-product raw material, NEC
5191	Wholesaling: Farm supplies
8699	Services: Membership organizations, NEC

[1]Standard Industrial Classification System 3- or 4-digit industry code.
NEC is "not elsewhere classified."

we defined as being "farm-related." These include agricultural services, farm suppliers, and most food processors relating to crops. Since they would likely be less affected by CRP, farm-related establishments devoted exclusively to livestock, such as meat processors and veterinary services, were excluded from this definition.

Beginning Farmer Model

In modeling the beginning farmer response to CRP enrollment, CRP's local importance was measured as the proportion of county cropland enrolled in CRP. Using this measure, we selected a group of high-CRP and matching counties which was different from the one used in the population and employment analysis. As a result, even though we used many of the same explanatory variables discussed above, the means of the high-CRP and matching counties differ slightly from those reported in tables A.1 and A.2. Nonetheless, for expositional ease they will not be reported.

The full list of explanatory variables considered for the beginning farmer models includes all of the demographic variables discussed above: the percent of population Black, Hispanic, Native American, under 18 years of age, or over 62 years of age. Many of the labor market and economic variables also enter the basic equation, in one form or another: the log of 1970-

80 population change; the log of 1980 population density; the percent of 1980 employment in agriculture, business services, manufacturing, mining, and recreation; percent of the civilian workforce employed in 1980; percent working outside the county; and median household income in 1979. Of the agricultural characteristics discussed earlier, the beginning farmer models included the proportion of sales going to very small farms (under $20,000 sales) and large farms (over $250,000 sales) and the proportion of farm operators working off-farm over 200 days a year.

Farm-sector variables not discussed earlier include the proportion of cropland in acreage reduction programs, the proportion of cropland not planted or diverted from production, the proportion of farm operators over 65 years of age, the number of farms in the county (which enters select models in log form), the proportion of farmland in crops, and the proportion of county land area devoted to farming. The basic equations also included the percentages of 1982 farm sales coming from specific commodities. For expositional ease, these data are not reported here. All farm-sector variables are from the 1982 Census of Agriculture. The dependent variables for the results reported in the text include the ratio of the number of young or short-tenure farm operators in 1997 relative to their numbers in 1982. These were further divided into ratios for each 5-year segment between 1982 and 1997. Identical models were estimated with the change in the share of all farmers that were young or short-tenure over this period as the dependent variable (not reported in the text for expositional ease). Descriptive statistics for each of these variables are reported in table A.5.

With over 35 possible explanatory variables, we used a backward stepwise regression procedure to narrow the set. Regressions were first estimated with the aggregate CRP-enrollment-to-cropland ratio. This variable was then replaced with similar ratios for whole- and for partial-farm acres.

Table A.5—Mean values of variables unique to the beginning farmer models

Variable description	Unit	Study counties	High-CRP[1]	Matched counties
Beginning farmer measures:				
Young farmers, 1997/1982	Percent	45.2	45.8	44.7
Short-tenure farmers, 1997/1982	Percent	69.7	74.0	73.4
CRP measures:				
CRP acres (1991-93)/cropland, 1982	Percent	8.0	26.8	4.8**
Whole-farm acres/cropland	Percent	3.0	11.1	1.6**
Farm and farm operator characteristics:				
Diverted acres/cropland, 1982	Percent	1.6	3.1	1.9**
Cropland not planted or diverted, 1982	Percent	29.4	30.3	29.2
Cropland/farmland, 1982	Percent	56.2	47.3	50.6**
Farmland/all county land, 1982	Percent	69.1	75.9	75.2
Number of farms, 1982	Number	720	476	667**

** indicates that the difference between high-CRP counties and their matched pairs is significantly greater than 0 at the 0.01 level.
[1]High-CRP counties have an average CRP acres-to-cropland ratio for 1991-93 exceeding 20 percent. Of the 1,481 study counties, 194 were high-CRP by this definition.

Source: 1982-1997 Census of Agriculture and the CRP Contracts file. In addition to the variables listed above, the models also included the proportion of total sales in 1982 from the full range of farm commodities (not reported for expositional ease).