Whose Votes Don't Count?:
An Analysis of Spoiled Ballots in the 2000 Florida Election

Philip A. Klinkner
Associate Professor of Government
Hamilton College
198 College Hill Road
Clinton, NY 13323
315-859-4344
pklinkne@hamilton.edu

This project began in response to media reports about the findings of the U.S. Commission on Civil Rights (USCCR) that indicated higher rates of ballot spoilage in Florida counties with larger numbers of blacks. I was intrigued by this result, so I decided to run my own independent analysis of the data. I should mention that I have no official connection to the USCCR, but I have met two of its members, Abigail Thernstrom and Professor Christopher Edley. I did not, however, contact them before undertaking this analysis. In addition, I did not know Professor Alan Lichtman who conducted the USCCR analysis. Since arriving at my findings, I have spoken by telephone with Professor Lichtman and Professor Edley to inform them of my conclusions. In summary, these findings are mine and mine alone.

As a first step, I obtained data on the dependent variable—the rate of spoiled ballots in each of Florida's counties. This information came from the Governor's Select Task Force On Election Procedures, Standards and Technology, conducted by the Collins Center for Public Policy (the report is available at http://www.collinscenter.org/info-url2660/info-url.htm).

The next step was to consider the different independent variables that might explain the differential rates of ballot spoilage. Among the list of possible suspects are the following:

Different types of voting systems:
As the Florida election controversy revealed, different types of voting systems have different rates of accuracy. Perhaps the differences in ballot spoilage rates could be explained by the fact that different Florida counties use different types of voting systems.

The type of voting system is indicated by four variables. **Op/P** refers to optical scan systems in which ballots are read at the precinct where the vote is cast. **Op/C** refers to optical scans systems in which ballots are collected from individual precincts and read at a central location. **Punchcard** refers to the now infamous punchcard voting systems. **Other** refers to the two

counties using different types of voting systems. One county uses the lever-machine system and the other uses paper/hand ballots. This information was obtained from the Governor's Select Task Force On Election Procedures, Standards and Technology, conducted by the Collins Center for Public Policy (the report is available at http://www.collinscenter.org/info-url2660/info-url.htm)

Turnout:
Turnout, defined as the percent of those registered who actually show up to vote, might influence turnout since it could mean more novice or inexperienced voters. High turnout might also create long lines at the polls and thus voters might be more concerned about completing their ballots quickly than with doing so accurately. Finally, high rates of turnout might also mean polling places in which the number of voters might swamp the available poll workers, thus making them less able to assist voters in completing ballots accurately or in tabulating votes accurately.

Turnout rates are the number of votes cast in the county divided by the number of registered voters in that county. Information on registered voters for each county is available from the Florida Elections Division website:
http://election.dos.state.fl.us/voterreg/vrArchive/2000voterreg.shtml#general

Data on the votes cast in each county is available at the Florida Elections Division website:
http://election.dos.state.fl.us/elections/resultsarchive/Index.asp

Gore %:
This is defined as the percent of the votes cast in the county for Al Gore. Perhaps the rate of spoiled ballots differed among Republicans and Democrats. Data on the presidential voter for each county is available at the Florida Elections Division website:
http://election.dos.state.fl.us/elections/resultsarchive/Index.asp

% Hispanic:
Spoiled ballots might be more common among Hispanics for a variety of reasons, namely less familiarity with English and that recent immigrants might have less knowledge about voting procedures and politics. This data was obtained from the 2000 U.S. Census available at
http://factfinder.census.gov/servlet/BasicFactsServlet

Median Income:
Spoiled ballots might be more common among poor people, and/or counties with low incomes might be less able to afford more accurate voting systems. This data was obtained from the 1990 U.S. Census at http://venus.census.gov/cdrom/lookup

Literacy:
Many have suggested that less literate voters might be more inclined to spoil their ballots since they will be less capable of reading and following instructions. Data on the literacy by county in Florida is from the 1992 National Adult Literacy Survey. The numbers indicate the percentage of adults in the county at Level 1 Literacy. This is the lowest level of literacy and persons at this level are unable to complete simple reading tasks, such as understanding a bus schedule. The data is available from the website of the Florida Literacy Coalition at
http://www.floridaliteracy.org/level_one.htm

Education:
Like literacy, low education levels might influence rates of ballot spoilage. For this I used the percent of persons aged 25 or older that have completed less than the 9th grade. This data was obtained from the 1990 U.S. Census at http://venus.census.gov/cdrom/lookup

% Black Registered Voters:
As the report of the U.S. Commission on Civil Rights claimed and as media reports after the election indicated, the rate of ballot spoilage seemed higher in largely black areas. For this I used the percent of registered voters in the county who are black. Information on registered voters by race for each county is available from the Florida Elections Division website: http://election.dos.state.fl.us/voterreg/vrArchive/2000voterreg.shtml#general

Voters per Precinct:
As with turnout, spoiled ballots might result from voters who have had to wait in line. This factor might be reflected in the number of voters per precinct within the county. In addition, with more voters per precinct, it might also be the case that there are fewer election workers to assist in accurately filling out ballots. The number of voters along with the number of precincts in each county is available at the Florida Elections Division website: http://election.dos.state.fl.us/elections/resultsarchive/Index.asp

Increase in Registration:
Spoiled ballots might result from increased numbers of first time voters. Since these voters are, by definition, less familiar with the process, they might be more likely to spoil their ballots. One indication of more first time voters might be increased numbers of registered voters over a previous year, in this case 1996, the year of the last presidential election. Information on registered voters for each county in 2000 is available from the Florida Elections Division website:
http://election.dos.state.fl.us/voterreg/vrArchive/2000voterreg.shtml#general

Information on registered voters in 1996 is available from the Florida Elections Division website: http://election.dos.state.fl.us/voterreg/vrArchive/1996voterreg.shtml#General

Increase in Voting:
Another indication of more first time voters might be an increase in the number of actual voters from one election to another. In this case, I've used the percentage increase in voters for each county from 1996 to 2000. Data on election results from both 1996 and 2000 is available from the Florida Elections Division website:
http://election.dos.state.fl.us/elections/resultsarchive/Index.asp

Other variables: In addition to these variables, I also ran models with the following variables, all of which proved either substantively and/or statistically insignificant:

--county crime rates,
--percent of elderly population,
--percent of population under 25,
--party of the county election supervisor,
--percent of population with less than a high school diploma,
--percent of population with some college education,

--percent of population in rural areas,
--percent of population English-only speakers,
--county population density,
--percent of blacks with less than 9th grade education,
--percent of blacks with less than a high school diploma,
--percent of blacks with some college education,
--increase in percent of registered voters who are black from 1996 to 2000

I then ran a regression model using the fourteen independent variables previously listed. The regression was run using SPSS 10.0 for the Macintosh. The results are as follows:

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.933(a)	.870	.837	1.257534117413E-02
a Predictors: (Constant), % Increase Vote 96-00, OTHER, Level 1 Literacy, Gore%, Punchcard, 1989 Median $, Turnout, % Hispanic, Voters/ Precincts, Opt/C, % Black Reg 2000, 96-00 % Increase Total Reg, % < 9th				

Coefficients(a)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.06236	.032		1.950	.057
	Opt/C	.04211	.005	.544	8.312	.000**
	Punchcard	.03469	.004	.541	8.612	.000**
	OTHER	.02278	.010	.127	2.379	.021*
	Turnout	-.06523	.031	-.129	-2.087	.042*
	Gore%	-.04972	.022	-.146	-2.309	.025*
	% Hispanic	-.003395	.024	-.011	-.143	.887
	1989 Median $	.0000004052	.000	.061	.749	.457
	Level 1 Literacy	.03147	.054	.058	.588	.559
	% < 9th	.05617	.055	.094	1.030	.308
	% Black Reg 2000	.133	.024	.392	5.442	.000**
	96-00 % Increase Total Reg	-.001337	.002	-.058	-.718	.476
	Voters/ Precincts	-.00001666	.000	-.187	-2.763	.008**
	% Increase Vote 96-00	-.01598	.020	-.044	-.781	.439
a Dependent Variable: % Spoiled *Significant at the 95% level **Significant at the 99% level						

Excluded Variables(b)

Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics	Tolerance
1	Opt/P	.(a)	.	.	.	.000	
a Predictors in the Model: (Constant), % Increase Vote 96-00, OTHER, Level 1 Literacy, Gore%, Punchcard, 1989 Median $, Turnout, % Hispanic, Voters/ Precincts, Opt/C, % Black Reg 2000, 96-00 % Increase Total Reg, % < 9th							
b Dependent Variable: % Spoiled							

As the model shows, the following variables were not significant:

1. % Hispanic
2. 1989 Median $
3. Level 1 Literacy[1]
4. % <9th
5. 96-00% Increase in Total Reg
6. % Increase Vote 96-00

Conversely, the following variables were statistically significant at the .05 level or greater.

1. Op/C
2. Op/P
3. Punchcard
4. Other
5. Turnout
6. Gore %
7. % Black Reg Voters
8. Voters/Precincts

Interestingly, education and income appear to have no effect on the rate of spoiled ballots. Thus there is little evidence in the data for the claim that spoiled ballots in Florida resulted mostly from the individual errors of voters who lacked the education or experience to cast accurate

[1] In his forthcoming book, Judge Richard Posner has his own analysis of Florida's spoiled ballots. According to the book's introduction (available on the web at: http://www.pupress.princeton.edu/titles/7118.html) Posner finds that literacy is a significant variable for explaining spoiled ballots. I checked Posner's literacy data however (available at: http://home.uchicago.edu/~rposner/election) and found that most of the data corresponded exactly with my own, indicating that we were working from the same data source--the 1992 National Adult Literacy Survey. Most of the data, however, is not all of the data, and I found that in Posner's data a handful of counties, all beginning with the letter M, were in error. Thus, Posner's assertion that literacy is a factor in ballot spoilage rests upon inaccurate data.

votes. This fact can be seen in the following regression, which I call the "stupid voter" model. In it, I use only those factors that are under the control of the voter: literacy, education, poverty (arguably under the control of individuals, but I'll put it in anyway), and first time voters (measured by the increase in registration from 1996 to 2000 and the increase in votes from 1996 to 2000).

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.710	.504	.452	2.307877377782E-02

a Predictors: (Constant), % in Poverty 1997, % Increase Vote 96-00, 96-00 % Increase Total Reg, % < 9th, Level 1 Literacy, % No HSG

Coefficients

		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
Model		B	Std. Error	Beta		
1	(Constant)	-4.682E-02	.018		-2.630	.011*
	Level 1 Literacy	.342	.096	.634	3.562	.001*
	% < 9th	.169	.173	.283	.976	.333
	% No HSG	-5.375E-02	.102	-.154	-.525	.602
	% Increase Vote 96-00	-2.460E-02	.034	-.068	-.731	.468
	96-00 % Increase Total Reg	-4.789E-03	.003	-.207	-1.814	.075
	% in Poverty 1997	2.459E-04	.001	.041	.256	.798

a Dependent Variable: % Spoiled
*Significant at the 95% level
**Significant at the 99% level

As the regressions for the "stupid voter" model indicates, these individually controlled variables exert little explanatory power. The adjusted r^2 is only .452, meaning that the model explains less than half the variance in the pattern of spoiled ballots across Florida counties. Furthermore, none of the variables, save for literacy is statistically significant. This model and the previous one show quite clear that the pattern of spoiled ballots in Florida was much more influenced by systemic factors rather than individual ones.

I then re-ran the model using only the statistically significant variables. The results are as follows:

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.931(a)	.866	.850	1.204969268270E-02
a Predictors: (Constant), Voters/ Precincts, Punchcard, OTHER, % Black Reg 2000, Turnout, Gore%, Opt/C				

Coefficients(a)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	0.09650	.018		5.381	.000**
	Opt/C	0.04394	.004	.592	10.137	.000**
	Punchcard	0.03465	.004	.537	9.838	.000**
	OTHER	0.02272	.009	.125	2.563	.013*
	Turnout	-0.08160	.027	-.159	-3.071	.003**
	Gore%	-0.04831	.019	-.141	-2.554	.013*
	% Black Reg 2000	0.13700	.019	.399	7.147	.000**
	Voters/ Precincts	-0.00002	.000	-.233	-4.172	.000**

a Dependent Variable: % Spoiled
*Significant at the 95% level
**Significant at the 99% level

Excluded Variables(b)

Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics	Tolerance
1	Opt/P	.(a)	.	.	.	.000	

a Predictors in the Model: (Constant), Voters/ Precincts, Punchcard, OTHER, % Black Reg 2000, Turnout, Gore%, Opt/C

b Dependent Variable: % Spoiled

The previous model seems to do a good job of explaining the rate of ballot spoilage in Florida counties. The adjusted r^2 is .85, indicating that 85 percent of the variation between counties in the rate of spoiled ballots can be explained by these variables.

The model shows that the type of voting system used in a county has a clear impact on the rate of ballot spoilage. In addition, so does the level of turnout, Gore's percent of the vote, the percent of registered voters who are black, and the number of voters per precincts. The fact that higher levels of turnout and more voters per precinct are negatively associated with ballot spoilage seems a bit counterintuitive. On the other hand, higher levels of turnout might also reflect greater political interest and knowledge, and thus probably mean less chance of spoiled or mistaken ballots. With the number of voters per precinct, the result is probably due to the fact that where counties have reduced the number of precincts, they will likely have more poll workers and election officials per precinct, thus making it easier for voters to obtain help in filling out their ballot.

The determine more precisely the impact of different voting systems on the other variables I set up several interactive variables by multiplying the following variable with one another:

Op/P * Turnout
Op/P * Gore%
Op/P * % Black Reg Voters
Op/P * Voters/Precincts
Op/C * Turnout
Op/C * Gore%
Op/C * % Black Reg Voters
Op/C * Voters/Precincts
Punch * Turnout
Punch * Gore%
Punch * % Black Reg Voters
Punch * Voters/Precincts
Other * Turnout
Other * Gore%
Other * % Black Reg Voters
Other * Voters/Precincts

I then re-ran the model using the existing variables and these new interactive variables. After dropping out the non-significant variables, I came up with the following results:

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.945(a)	.892	.875	1.100907477018E-02
a Predictors: (Constant), Punch*% Gore, Voters/ Precincts, % Black Reg 2000, Turnout, Gore%, Opt/P, Op/P*Voters/Precincts, Op/P*Gore%, Punchcard				

Coefficients(a)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	0.08397	.024		3.568	.001**
	Opt/P	-0.01099	.020	-.173	-.546	.587
	Punchcard	0.06845	.021	1.061	3.184	.002**
	Turnout	-0.06101	.025	-.119	-2.429	.018*
	Gore%	0.08671	.045	.253	1.910	.061
	% Black Reg 2000	0.11400	.019	.334	5.938	.000**

	Voters/ Precincts	-0.00004	.000	-.446	-6.339	.000**
	Op/P*Gore%	-0.13600	.049	-.918	-2.769	.008**
	Op/P*Voters/Precincts	0.00003	.000	.523	3.415	.001**
	Punch*% Gore	-0.17100	.050	-1.222	-3.448	.001**
a Dependent Variable: % Spoiled *Significant at the 95% level **Significant at the 99% level						

Thus, the statistically significant variables that exercise an independent effect are:

1. Punchcard
2. Turnout
3. % of Black Reg Voters
4. Voters/Precincts
5. Op/P * Gore%
6. Op/P * Voters/Precincts
7. Punch * % Gore

In overall terms, the following factors led to increased levels of spoiled ballots:

1. Counties with punchcard ballots
2. Counties with higher percentages of black registered voters
3. Counties with Op/P voting systems with higher numbers of voters per precinct

In addition, the following factors led to lower levels of spoiled ballots:

1. Counties with higher levels of turnout
2. Counties with more voters per precinct
3. Counties with Op/P voting systems with higher percentages for Gore.
4. Counties with punchcard voting systems with higher percentages for Gore.

One aspect of these regressions seems incongruous. According to initial model, the percent for Gore negatively correlates with the percent of spoiled ballots. In other words, if a county has a higher percent for Gore, it is likely to have fewer spoiled ballots. On the other hand, the percent of registered voters who are black is positively correlated with spoiled ballots—counties with a greater percentage of black registered voters were likely to have more spoiled ballots. This finding seems odd, since the percent of black voters and the percent for Gore are moderately correlated (.415). Another way to look at it is that the model seems to suggest that ballot spoilage is likely to be highest in counties with high percentages of blacks and low votes for Gore.

This is an interesting finding, since if we were to suspect racial disenfranchisement in Florida, we would expect to find it in certain types of counties. Racial disenfranchisement would be least likely to take place in strongly Democratic counties. Election officials would have no incentive to disenfranchise some of their most loyal voters. Conversely, we would also be unlikely to find racial disenfranchisement in heavily Republican areas with very few black voters. Such tactics would yield little benefit and most likely they would be difficult to carry, since the black vote would be less concentrated and identifiable.

On the other hand, racial disenfranchisement would be most likely in counties where there is a significant or at least non-trivial percentage of black voters, but at the same time the county is strongly Republican. In such cases there are enough black votes to create political incentives for racial disenfranchisement and the black vote would be more concentrated and identifiable.

To test this possibility, I developed an additional interactive variable, the percent of registered voters who are black multiplied by the winning vote margin for George W. Bush. This variable provides a good proxy for the types of counties where racial disenfranchisement is most likely to occur. Counties that rank highest have a sizeable black vote and a larger Republican vote. The lowest ranking counties have both a large black vote and go strongly Democratic.

I then re-ran the regressions including this variable.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.972	.945	.927	8.411314622509E-03

a Predictors: (Constant), Other*BushMargin, Gore%, Opt/C, Turnout, Other*% Black Reg 2000, Punch*% Black Reg 2000, Punch*BushMargin, Op/C*BushMargin, Voters/ Precincts, Punch*Voters/Precincts, Op/P*BushMargin, % Black Reg 2000, BushMargin*%Black Reg 2000, Op/P*Gore%, Punchcard, Op/C*Gore%

Coefficients

		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
Model		B	Std. Error	Beta		
1	(Constant)	7.885E-02	.016		4.946	.000**
	Opt/C	7.217E-03	.019	.097	.372	.711
	Punchcard	.109	.020	1.687	5.486	.000**
	Turnout	-6.541E-02	.020	-.128	-3.306	.002**
	Gore%	-.185	.039	-.540	-4.757	.000**
	% Black Reg 2000	7.116E-02	.022	.208	3.287	.002**
	Voters/ Precincts	-7.628E-06	.000	-.087	-1.476	.146
	Op/C*Gore%	.256	.056	1.478	4.551	.000**
	Op/P*Gore%	.139	.048	.939	2.878	.006**
	Punch*% Black Reg2000	8.820E-02	.032	.206	2.749	.008**
	Punch*Voters/Precincts	-2.264E-05	.000	-.351	-2.899	.006**
	Other*% Black Reg 2000	.662	.155	.318	4.280	.000**
	BushMargin*%Black Reg 2000	4.784E-06	.000	.663	5.580	.000**
	Op/C*BushMargin	-2.650E-06	.000	-.169	-3.643	.001**
	Op/P*BushMargin	-5.214E-07	.000	-.158	-2.261	.028*
	Punch*BushMargin	-7.523E-07	.000	-.747	-5.588	.000**

	Other*BushMargin	6.536E-06	.000	.189	2.342	.023*

a Dependent Variable: % Spoiled
*Significant at the 95% level
**Significant at the 99% level

Excluded Variables

		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics
Model						Tolerance
1	Opt/P	.	.	.	.	.000
	OTHER	.	.	.	.	.000
	Bush Margin	.	.	.	.	.000
	Punch*% Gore	.	.	.	.	.000
	Other*Turnout	.	.	.	.	.000
	Other*% Gore	.	.	.	.	.000
	Other*Voters/Precincts	.	.	.	.	.000
	Other*BushMargin%Black	.	.	.	.	.000

a Predictors in the Model: (Constant), Other*BushMargin, Gore%, Opt/C, Turnout, Other*% Black Reg 2000, Punch*% Black Reg 2000, Punch*BushMargin, Op/C*BushMargin, Voters/ Precincts, Punch*Voters/Precincts, Op/P*BushMargin, % Black Reg 2000, BushMargin*%Black Reg 2000, Op/P*Gore%, Punchcard, Op/C*Gore%
b Dependent Variable: % Spoiled

As the tables indicate, the adjusted r^2 is now quite high—the model explains over 92 percent of the variance in spoiled ballots. In addition, the percent of black voters remains significant. Finally, in areas where the combined result of multiplying the percent of voters who are black by the voter margin for Bush is positive, there is a positive correlation with spoiled ballots. To put it another way, not only does being black matter in the model, it also matters where you are black. Strongly Republican areas that also had a sizeable proportion of blacks had a greater incidence of spoiled ballots. While this finding is only suggestive, it is exactly what one would expect to find in a situation where racial disenfranchisement is likely to occur--black voters are a sizeable part of the electorate, but lacked the political power to ensure that their ballots are counted accurately and fairly.

In conclusion, this analysis offers two important findings:

1. There is no evidence that higher rates of spoiled ballots resulted from such individual factors as education and literacy. Instead, the factors influencing spoiled ballots were systemic. Thus, rather than speaking of individuals who spoiled their ballots, we should speak of individuals who were placed in situations in which it was more likely that their ballots would be spoiled. Furthermore, this finding indicates that any effort to reduce the rate of spoiled ballots must focus on systemic solutions--improved technology, more and better election workers, and stronger efforts to investigate and prosecute any instances of corruption and/or racial disenfranchisement.

2. Even after controlling for other factors, rates of ballot spoilage remain higher in predominantly black areas than in other areas of Florida. As the last model indicates, with all else being equal, for every 1-point increase in the percentage of registered voters who are black, there was a .07 percentage point increase in spoiled ballots.

In addition, these rates were even higher where substantial numbers of blacks were found in counties with large margins for George W. Bush. All of this corresponds to and further reinforces the findings of the USCCR that there is evidence of racial disenfranchisement in the 2000 election in Florida. Consequently, it is important that federal authorities should investigate this matter more thoroughly.

Addendum 1: Evaluating John Lott's Analysis

In response to the findings of the majority of USCCR, the minority members of the Commission submitted an alternative analysis vote spoilage in Florida by Professor John Lott (his report is available at http://www.manhattan-institute.org). According to Professor Lott's analysis, the percentage of black voters in a county has no statistical relationship to the percent of spoiled ballots in that county. Rather, the percent of people in poverty is a more important factor for explaining the rate of ballot spoilage. In this section I will evaluate Professor Lott's analysis.

Lott begins his analysis by suggesting that Lichtman's cross-sectional analysis is insufficient, arguing that if African Americans were more likely to spoil their ballots, then changes in spoiled ballots across time should closely correlate with changes in the percentage of African Americans across the same period of time. In a series of scatterplots, Lott shows that there is little if any relationship between the change in percent of spoiled ballots between 1996 and the change in percent of voters who are black between the same years. But Lott makes a critical error by assuming that all other factors that might influence ballot spoilage remained equal between 1996 and 2000. This is extremely doubtful. Consequently, even if increased percentages of black voters led to increased percentages of spoiled ballots for a particular county, this finding might not be apparent if, for example, that county moved to a more accurate voting system. Evidence of this can be seen in Table 1.

Table 1.

Year 1		Black Spoilage Rate	10%	Black Voters	1000	Black Spoiled Ballots	100
		White Spoilage Rate	5%	White Voters	1000	White Spoiled Ballots	50
				Total Voters	2000	Total Spoiled Ballots	150
				% Black	50%	% Spoiled	7.50%
Year 2a	(Increased % Black, Same Voting System)						
		Black Spoilage Rate	10%	Black Voters	1200	Black Spoiled Ballots	120
		White Spoilage Rate	5%	White Voters	1000	White Spoiled Ballots	50
				Total Voters	2200	Total Spoiled Ballots	170
				% Black	55%	% Spoiled	7.70%

Year 2b	(Increased % Blacks, More Accurate Voting System)					
	Black Spoilage Rate	5%	Black Voters	1200	Black Spoiled Ballots	60
	White Spoilage Rate	2.5%	White Voters	1000	White Spoiled Ballots	25
			Total Voters	2200	Total Spoiled Ballots	85
			% Black	55%	% Spoiled	4%

In each year, blacks spoiled their ballots at twice the rate of whites. In Year 1, blacks made up 50 percent of the county population and the rate of spoiled ballots was 7.5 percent. Year 2a represents what Lott expects to find. The percent of voters who are black rises from 50 to 55 percent, leading to a corresponding increase in the percent of spoiled ballots from 7.5 percent to 7.7 percent. But in Year 2a, nothing changes but the percent of voters who are black.

In Year 2b, however, two things change. In addition to the increase in the percent of voters who are black from 50 to 55 percent (as in Year 2a), the county also institutes a new voting systems, cutting the spoilage rate in half, from 10 to 5 percent for blacks and from 5 to 2.5 percent for whites. This change means that in Year 2b, the percent of spoiled ballots declines to 4 percent from 7.5 percent in Year 1, despite the fact that blacks now make up a larger share of the county's voters. Consequently, it is impossible to know the true relationship between change in the percent of voters who are black and change in the percent of spoiled ballots without controlling for other factors, especially voting systems.

Lott then goes on to develop a series of models that, he argues, better explain the rate of spoiled ballots in Florida counties. According to the results of his models, the percentage of black voters is not a significant factor in explaining the percentage of spoiled ballots in a county. His models also suggest a county's poverty rate has more to do with spoiled ballots than the racial and ethnic makeup of the county.

The ultimate test of any regression model is the amount of variance in the dependent variable that is explained by the independent variables. In this case, how much of the variation in the percent of spoiled ballots across counties is explained. On this score Lott's models are inferior to those that I have developed. Of the eight models listed in Lott's Table 2, the highest r^2 is .7859, meaning that his independent variables explain almost 79 percent of the variation in spoiled ballots across Florida counties (I'll give Professor Lott the benefit of the doubt by assuming that these are adjusted r^2's. If not, then his models have even less explanatory power). Furthermore, a model using only variables for the different types of voting systems for each county yields an adjusted r^2 of .594. This indicates that the bulk of the explanatory power in Lott's models comes from these variables.

In comparison, even the simplest model that I developed has an adjusted r^2 of .837 and my final model had an r^2 of .927. Put in other terms, my final model is 17 percent more powerful in explaining the pattern of spoiled ballots than Lott's. Furthermore, I also entered several of Lott's

variables, namely party of the county election supervisor, percent Hispanic population, and county median income into my earlier model and all came up as statistically insignificant.

Nonetheless, perhaps Lott is on to something with his variables and they might turn up as significant in my final model. Consequently, I reran my model and added in the variables for median income, percent of the population in poverty, and variables for the party of the county election supervisor.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.972	.946	.922	8.698121304546E-03

a Predictors: (Constant), NP Supervisor, Bush Margin, % in Poverty 1997, OTHER, Op/C*BushMargin, Punchcard, Turnout, Op/P*BushMargin, % Black Reg 2000, Op/C*Gore%, R Supervisor, Voters/ Precincts, Other*BushMargin, Gore%, Punch*% Black Reg 2000, 1997 Median Household Income, Punch*Voters/Precincts, BushMargin*%Black Reg 2000, Opt/C, Op/P*Gore%

Coefficients

		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
Model		B	Std. Error	Beta		
1	(Constant)	7.706E-02	.032		2.407	.020*
	Opt/C	5.370E-03	.021	.072	.259	.797
	Punchcard	.107	.021	1.655	5.139	.000**
	OTHER	8.580E-02	.022	.472	3.904	.000**
	Turnout	-6.840E-02	.021	-.133	-3.278	.002**
	Gore%	-.181	.042	-.530	-4.335	.000**
	Bush Margin	-7.507E-07	.000	-.788	-5.147	.000**
	% Black Reg 2000	7.136E-02	.027	.209	2.641	.011*
	Voters/ Precincts	-8.998E-06	.000	-.103	-1.516	.136
	Op/C*Gore%	.256	.059	1.479	4.329	.000**
	Op/P*Gore%	.133	.051	.901	2.602	.012*
	Punch*% Black Reg 2000	8.819E-02	.035	.206	2.538	.015*
	Punch*Voters/Precincts	-2.296E-05	.000	-.355	-2.773	.008**
	BushMargin*%Black Reg 2000	4.854E-06	.000	.673	5.123	.000**
	Op/C*BushMargin	-1.941E-06	.000	-.124	-2.451	.018*
	Op/P*BushMargin	1.884E-07	.000	.057	.826	.413
	Other*BushMargin	-2.455E-06	.000	-.071	-1.088	.282
	% in Poverty 1997	-3.233E-05	.001	-.005	-.051	.959
	1997 Median	2.115E-07	.000	.037	.443	.660

	Household Income					
	R Supervisor	-1.617E-05	.004	.000	-.004	.997
	NP Supervisor	-2.651E-04	.005	-.002	-.052	.959

a Dependent Variable: % Spoiled
*Significant at the 95% level
**Significant at the 99% level

Excluded Variables

		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics
Model						Tolerance
1	Opt/P	.	.	.	.	.000
	Punch*% Gore	.	.	.	.	.000
	Other*Turnout	.	.	.	.	.000
	Other*% Gore	.	.	.	.	.000
	Other*% Black Reg 2000	.	.	.	.	.000
	Other*Voters/Precincts	.	.	.	.	.000
	Punch*BushMargin	.	.	.	.	.000
	Other*BushMargin%Black	.	.	.	.	.000
	D Supervisor	.	.	.	.	.000

a Predictors in the Model: (Constant), NP Supervisor, Bush Margin, % in Poverty 1997, OTHER, Op/C*BushMargin, Punchcard, Turnout, Op/P*BushMargin, % Black Reg 2000, Op/C*Gore%, R Supervisor, Voters/ Precincts, Other*BushMargin, Gore%, Punch*% Black Reg 2000, 1997 Median Household Income, Punch*Voters/Precincts, BushMargin*%Black Reg 2000, Opt/C, Op/P*Gore%
b Dependent Variable: % Spoiled

As the tables indicate, adding in Lott's variables to my final model increases the unadjusted r^2 only from .945 to .946, a trivial increase. Indeed, I added a random variable (the number of letters in the name of the county) and the r^2 also increased from .945 to .946, indicating that the addition of Lott's variables adds no more explanatory power than the addition of a random variable. (In fact, the significance of this random variable was .283, making it more statistically significant that any of Lott's variables!)

Since the addition of any variable (even a random variable like the number of letters in the name of the county) will increase the r^2 of a model, the more important statistic is the adjusted r^2, which controls for the number of variables in the model. Adding Lott's variables to my model drops the adjusted r^2 from .927 to .922. In other words, adding in the variables that Lott claims have the most explanatory power actually makes my model less, not more, powerful. Furthermore, none of Lott's variables is statistically significant, and their addition to the model causes only two of the original variables, Op/P*BushMargin and Other*BushMargin, to fall out of significance. The percent of registered voters who are black remains statistically significant and the correlation coefficient remains largely the same as in my earlier model.

In conclusion, Lott's findings do not hold up under scrutiny. Not only do they under-explain the variance in the rate of spoiled ballots, but when his variables are added to a more sophisticated model, they lack statistical significance. As a result, nothing in Lott's analysis detracts from the finding of the USCCR majority report or the analysis that I've offered here.

Addendum 2: First Time Voters

In the earlier models I used two proxies for first time voters: increase in registration from 1996 to 2000 and increase in the number of votes cast from 1996 to 2000. Neither variable was significant. Fortunately, I was just able to come up with the number of first time voters from the Florida Secretary of State Voter File (available on CD-ROM for free by calling the Florida Elections Division). These files contain data on every current voter in Florida, including their voting history since 1994. Using this data, I determined how many voters in each county voted only in the 2000 general election. While this would also included some non-first time voters (voters who voted prior to 1994 or voters who voted in other states by voted in Florida for the first time in 2000), it is a better estimate than the previous ones.

I reran my simple model and included a new variable for the % of first time voters. The results are as follows:

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.934	.873	.855	1.161578611804E-02

a Predictors: (Constant), %First Timers, OTHER, Turnout, Punchcard, Gore%, % Black Reg 2000, Opt/C, Voters/ Precincts

Coefficients

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	9.210E-02	.019		4.831	.000**
	Opt/C	4.462E-02	.004	.621	10.191	.000**
	Punchcard	3.374E-02	.003	.533	9.679	.000**
	OTHER	2.277E-02	.009	.130	2.619	.011*
	Turnout	-7.599E-02	.027	-.148	-2.862	.006**
	Gore%	-4.906E-02	.019	-.149	-2.552	.013*
	% Black Reg 2000	.139	.020	.419	7.094	.000**
	Voters/ Precincts	-1.697E-05	.000	-.195	-3.046	.004**
	%First Timers	-1.360E-02	.037	-.025	-.365	.716

a Dependent Variable: % Spoiled
*Significant at the 95% level
**Significant at the 99% level

Excluded Variables

	Beta In	t	Sig.	Partial Correlation	Collinearity Statistics

Model						Tolerance
1	Opt/P	.	.	.	.	.000

a Predictors in the Model: (Constant), %First Timers, OTHER, Turnout, Punchcard, Gore%, % Black Reg 2000, Opt/C, Voters/ Precincts
b Dependent Variable: % Spoiled

As these results show, the variable for first time voters is not statistically significant. Moreover, it negatively correlated with spoiled ballots—meaning that counties with more first time voters had a higher, not lower, rate of spoiled ballots. All of this is further evidence of the soundness of my models and of the lack of statistical support for the notion that rates of spoiled ballots stemmed from individual-level factors.