Kenneth Piercy
(kpiercy@u.washington.edu)
Economics 482, Professor Startz
Research Paper, Final Draft

Executive Summary

The frequency of abortions varies substantially across counties in Washington State, both in relation to the number of childbirths that occur (the abortion ratio) and the number of childbearing women that live in each county (the abortion rate). The purpose of this paper is to investigate what factors influence the occurrence of abortion, in order to provide objective information that will prove useful in the discussion and analysis of such a sensitive and politically charged topic. This study attempts to explain these cross county differences via the method of least squares regression.

Two regressions are presented, using data for each of Washington's 39 counties from the years 1990 and 2000. The dependent variables are the abortion rate and the abortion ratio. Independent variables are based on education, income, religious adherence, demographic factors, poverty levels, and racial characteristics of each county.

This study contains a greater number of significant coefficients for the abortion ratio test, which implies that the variables studied have a stronger influence on the choice between carrying a pregnancy to term or having an abortion than on the total rate of abortions per woman of childbearing age. Increases in the percentage of the population who claim to be religious adherents, the percentage of the population of Asian descent, and the percentage of households that are families all lead to a lower abortion ratio. Real income and the percentage of a county that is of African-American descent lead to an increase in the abortion ratio. Education level is linked to increases in the abortion ratio, and a decrease in the abortion rate, but these values are not statistically significant at the five percent level.

The study does not incorporate "supply side" factors (such as cost and provider availability), as these data are not publicly available in a method conducive to statistical research. The implicit assumption is therefore that women of differing counties face the same opportunities to access of abortion providers. An obvious (and necessary) point of potential improvement in this study would involve a detailed study of access to services over time if one is interested in gauging the efficacy and potential location of abortion providers in the state of Washington.

Kenneth Piercy
(kpiercy@u.washington.edu)
Economics 482, Professor Startz
Research Paper, Final Draft

Factors Affecting the Demand for Abortions in Washington State

Subject and Sources of Data

According to data released by the Washington State department of health for the years 2000 and 1990, the frequency of abortions varies greatly across Washington State. The health department data classify abortion frequency in two ways: the abortion ratio, defined as the number of abortions per live birth, and the abortion rate, which is the number of abortions in each county per 1000 women of child bearing age (between the ages of 15 and 44). The abortion rate varies from a low of 6.8 in Stevens county to a high of 24.7 and 24.8 in Pierce and King counties for the year 2000. The abortion ratio for 2000 is highest in San Juan county, with a value of 533, and lowest in Grant County, which has a value of 101 (please see appendix 1 for a detailed presentation of these data).

The goal of this paper is to determine what factors influence the frequency of abortion, using data on abortions, income, and other social and demographic variables in each county. Regardless of where one stands in the politically and religiously charged issue of abortion rights and access, having a statistically accurate representation of these influencing factors could prove quite useful. From a "pro-life" viewpoint, the results could be beneficial in deciding what social groups should be targeted for adoption and abstinence counseling. "Pro-choice" individuals could utilize this data to make focused decisions on where to offer abortion services, and where demand for these services will be the greatest. Optimally, a linear least squares regression will offer great insight into how specific social and economic variables affect abortions performed in Washington state.

This study regresses the abortion rate and ratio, on economic, racial, and demographic indicators for the 39 counties of Washington State using 1990 and 2000 data. Most of this data is from the U.S. Census Bureau (www.census.gov), the exceptions being the abortion data, which is from the State Department of Health (2000 data is available on their website, http://www.doh.wa.gov/ehsphl/chs/chs-data/abortion/viewdown.htm, and 1990 data is from the book *Selected pregnancy and induced abortion statistics, Washington State*), and the data used to measure religious participation, which is based on survey results from the American Religion Data Archive (www.thearda.com).

Dependent and Independent Variables

The dependent variables used in the regressions are the abortion rate (*rate* in EViews), which is the number of abortions performed per 1000 women in the county who are of childbearing age (between 15 and 44 years of age), and the abortion ratio, (*ratio)*, the number of abortions per 1000 live births.

The independent variables used in this regression can be sorted into the following subsections:

Race: Using county specific data from the census, racial characteristics of each county are measured in percentage (integer) terms, with the EViews variable names listed here in parenthesis: African-American (*afram*) Asian-American (*asian*) and Native-American (*nativeam*). Due to the somewhat confusing presentation of the Latin American and Latino population in each county, this group is omitted from the paper.

Income and economic status: The hypothesis that income has an effect on the number of abortions was one of the key motivations for this study. The value of the income coefficient could be positive (perhaps due to abortion being more accepted in high income society), or negative (perhaps due to the increase in the relative burden of the cost of child raising for lower income Washington residents). Income and economic status are captured in the regression via two variables: the percentage of the population with income below the poverty line in each county (*belowpov*) and the median income for a resident of each county (*realincome*). Median income figures for 1990 and 2000 were deflated via the consumer price index (www.bls.gov/cpi/), and are expressed in constant dollars with a base year of 1984.

Educational attainment: The level of education, which is often closely linked to economic status, is hypothesized to have an influence on the frequency of abortions. The regression uses the percent of the population above 25 years of age that has attained at least a high school diploma (*high*) and the percentage of the population between the ages of 25 and 34 that has at least college bachelor's degree *(bach*) in each county to incorporate the effects of education on abortion frequency. There is an obvious overlap in these variables, as a high school diploma (or its equivalent) is required to attain a bachelor's degree.

Religion: Differences in religious participation across counties are incorporated into the model via data from the American Religion Data Archive. Their study lists the number of avowed religious adherents in each county (including those of the Jewish, Christian, and Muslim faiths). These values are divided by the population of each county and the result multiplied by 100 to give them the same percent format as the other variables used in the model (listed as *adherents* in the regression output).

Household Status: The percentage of households that are a family unit is incorporated in this paper by the variable *families*, and the percentage of households that are a married family is included as a separate class (*married*). Distinguishing married families from the general family class allows the potential effects of marriage as a social institution on the abortion levels to be observed. The source for this material is the Census Bureau.

Access to Facilities: The preliminary draft of this paper included the author's estimation of the availability of abortion facilities in each county for the year 2000 as a dummy variable. The method was imprecise, and the state department of health does not provide easily accessible documentation of the number and location of abortion providers. Any estimates of 1990 provider status would be specious, and so the final version of the study leaves this variable out, implicitly assuming that access to providers is constant over the ten-year period.

Time variables (and related Issues): The intercept term is allowed to vary across the two time periods studied, enabling a more accurate estimation of the coefficients for each independent variable. The results of a chow test imply that the null hypothesis (the hypothesis that the difference in the coefficients for the years 1990 and 2000 is zero) must be rejected for the abortion ratio. All variables were thus regressed with interaction terms; *realincome* and *bach* had statistically significant (at 5%) differences over the two time periods. (The Chow test results are provided in appendix 4). Statistically significant interaction terms are kept in the final regression equations, as presented below.

The regression Model:

The models used for the abortion rate (*rate)* and the abortion ratio (*ratio*) are the following:

$$ratio = \beta0 + \beta1\,\text{NINETY} + \beta2 * \text{ADHERENTS} + \beta3 * \text{REALINCOME} + \beta4 * \text{MARRIED} + \beta5 * \text{FAMILIES} + \beta6 * \text{HIGH} + \beta7 * \text{BACH} + \beta8 * \text{AFRAM} + \beta9 * \text{ASIAN} + \beta10 * \text{NATIVEAM} + \beta11 * \text{BELOWPOV} + \beta12 * \text{BACH*NINETY} + B13 * \text{REALINCOME*NINETY}$$

$$rate = \beta0 + \beta1\,\text{NINETY} + \beta2 * \text{ADHERENTS} + \beta3 * \text{REALINCOME} + \beta4 * \text{MARRIED} + \beta5 * \text{FAMILIES} + \beta6 * \text{HIGH} + \beta7 * \text{BACH} + \beta8 * \text{AFRAM} + \beta9 * \text{ASIAN} + \beta10 * \text{NATIVEAM} + \beta11 * \text{BELOWPOV} + \beta12 * \text{BACH*NINETY}$$

Interpreting the Regression Results:

This study offers some interesting insights into factors that influence the abortion ratio and rate, though they must be viewed with a degree of skepticism. The results of each regression are presented independently.

The abortion ratio: The abortion ratio is measured as the number of abortions per 1000 live births. It therefore considers only women who are sexually active and pregnant, avoiding a potential problem with the abortion rate variable. The variables that are statistically significant at the 95% level with 63 degrees of freedom and have positive coefficient values are: *ninety* and (*1-ninety)* (*the* constant terms for the two sets of observations), *afram*, the percentage of the population of African-American descent, realincome, the median income for each county in cpi deflated dollars, and *bach**ninety*, the separate coefficient value for college graduates according to 1990 data. Statistically significant variables with a negative coefficient values are *families*, the percentage of households that are considered a family, *adherents*, the percentage of the population that claim religious faith, and *asian*, the percentage of the population of Asian descent.

The remaining variables are not statistically significant at the five percent level. An F-test for these remaining seven variables yields a probability value of 0.4250, which implies that we *cannot* reject the null hypothesis that these variables all have a value of zero. The R^2 value for this regression is 0.77935. These results indicate that the number of abortions, in relation to the number of pregnancies carried to term, is influenced by a county's average income, racial composition, and proportion of religious adherents.

The abortion Rate: The abortion rate is the number of abortions per 1000 women of childbearing age. This includes many women who cannot or will not become pregnant (such as those who use contraception or abstain from intercourse). Thus the rate is a much broader measure of abortions; it can be considered a per-capita measure of the population who could undergo this procedure. The regression results for the abortion rate are quite similar to those for the ratio, but with fewer statistically significant coefficients. At the 5% level, variables with positive coefficients are *ninety, (1-ninety)* (the value of the slope coefficients), *afram*, and *bach**ninety*. There are no statistically significant variables at the 5% level with a negative valued coefficient.

Table 1: abortion ratio and rate regression coefficients (sign of coefficient, significant at 5% level?):

Regression	Ninety	(1-ninety)	afram	asian	nativeam	realincome	realincome*ninety
Ratio	(+), yes	(+), yes	(+), yes	(-), no	(+), no	(+), Yes	(-), no
Rate	(+), yes	(+), yes	(+) yes	(-), no	(+), no	(+), no	n/a

	adherents	families	belowpov	high	bach	bach*ninety	Married
Ratio	(-), Yes	(-), yes	(-), no	(-), no	(-) no	(+), yes	(+), no
Rate	(-), no	(-), no	(-), no	(-), no	(-), no	(+), yes	(-), no

<u>Conclusion and Caveats:</u>

The results of this study offer a statistical affirmation of general intuition and a few valuable insights. Many religions prohibit the practice of abortions among adherents and promote abstinence until marriage, and a higher percent of religious adherents in a county therefore would likely lead to a decrease in abortions, whether viewed in relation to pregnancies or to the childbearing population as a whole. Married couples are often in stable relationships and thus may have a greater ability to handle unexpected pregnancy, and therefore may defer abortions in favor of child raising. These intuitions are reinforced by the results of this study. Increases in per-county median income imply increases in both the abortion rate and ratio. Though the percentage of African-American residents is quite small in all Washington counties, there is a strong correlation between this percentage and increases in abortion statistics. Though not statistically significant at the five percent level, income variables (measured by median income and percentage below the poverty level) imply an increase in both the abortion rate and ratio. There are ample grounds for conjecture in these results; the study offers no concrete explanations for *why* these relationships occur.

A degree of caution must be exercised in interpreting these results. From a fundamental standpoint, the two regressions presented take aggregate values (such as income, religious participation, and the like) from the entire population of a county, while the incidence of abortion is confined to women of childbearing age. Social demographics do have an influence on whether a woman is sexually active, uses contraception, or has an abortion. However, a more revealing study would use specific observations for women who have been pregnant and had abortions, and not the aggregate data provided by the State Department of Health.

The accuracy of the data is also a possible point of contention in this study. Census data is provided in different forms for the years 1990 and 2000, and the author was required to calculate some percent variables (such as educational level) fromcensus reports for 1990. The values used by the author and the census may differ, leading to potential problems of heteroskedasticity and errors-in-variables. (Heteroskedasticity tests are presented in the appendix.)

Finally, there is a question of omitted variables that must be addressed. Abortion providers and costs (the "supply side" of this event) are omitted from the study due to insufficient and unavailable data. Addressing the availability of abortion services would provide greater insight into quantifying the occurrence of this medical practice.

Appendices:

1. Alphabetical List of Washington State counties, their abortion rates, and abortion ratios.

County name	*Abortion rate*		*Abortion ratio*	
	1990	*2000*	*1990*	*2000*
Adams	7	11.0	73	108
Asotin	14.0	10.4	196	171
Benton	13.0	14.8	175	215
Chelan	18	10.7	200	144
Clallam	20	16.0	278	272
Clark	13	16.7	187	233
Columbia	9	8.6	135	176
Cowlitz	10	15.0	136	217
Douglas	11	9.0	139	132
Ferry	13	9.1	200	174
Franklin	15	15.6	129	146
Garfield	8	--*	143	--*
Grant	13	9.6	146	101
Grays Harbor	25	18.5	347	301
Island	21	17.0	252	248
Jefferson	17	21.3	307	408
King	34	24.8	577	447
Kitsap	25	17.3	317	267
Kittitas	16	15.3	341	345
Klickitat	6	10.7	86	150
Lewis	17	14.8	229	222
Lincoln	12	10.4	196	182
Mason	20	19.5	288	304
Okanogan	15	13.5	177	200
Pacific	15	15.7	223	262
Pend Oreille	15	13.9	196	264
Pierce	32	24.7	418	383
San Juan	38	22.8	709	533
Skagit	20	16.0	275	235
Skamania	6	10.1	97	172
Snohomish	28	19.7	379	314
Spokane	21	16.0	325	258
Stevens	9	6.8	149	115
Thurston	26	21.3	429	378
Wahkiakum	5	14.7	115	321
Walla Walla	14	11.5	220	195
Whatcom	21	17.0	364	314
Whitman	18	10.8	520	311
Yakima	17	16.6	175	179

2. EViews regression output and heteroskedasticity tests.

The abortion rate:

Dependent Variable: RATE
Method: Least Squares
Date: 12/10/02 Time: 10:26
Sample(adjusted): 1 78
Included observations: 77
Excluded observations: 1 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
NINETY	50.98319	22.49236	2.266690	0.0268
1-NINETY	52.25486	22.54964	2.317326	0.0237
ADHERENTS	-0.084240	0.061229	-1.375823	0.1737
AFRAM	1.674175	0.633971	2.640777	0.0104
ASIAN	-0.533672	0.637955	-0.836536	0.4060
BACH	-0.090255	0.132997	-0.678627	0.4998
BACH*NINETY	0.424859	0.126669	3.354077	0.0013
BELOWPOV	-0.402001	0.308330	-1.303801	0.1970
FAMILIES	-0.036655	0.466833	-0.078519	0.9377
HIGH	-0.130025	0.157003	-0.828167	0.4107
MARRIED	-0.513638	0.382151	-1.344071	0.1837
REALINCOME	0.000519	0.000361	1.437788	0.1554
NATIVEAM	0.080666	0.192764	0.418469	0.6770

R-squared	0.630492	Mean dependent var	15.89695
Adjusted R-squared	0.561209	S.D. dependent var	6.394283
S.E. of regression	4.235655	Akaike info criterion	5.877693
Sum squared resid	1148.210	Schwarz criterion	6.273401
Log likelihood	-213.2912	Durbin-Watson stat	1.770094

Heteroskedasticity for the abortion rate:

White Heteroskedasticity Test:

F-statistic	3.135296	Probability	0.000303
Obs*R-squared	44.38118	Probability	0.004729

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 12/10/02 Time: 11:05
Sample: 1 78
Included observations: 77
Excluded observations: 1

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	99.79237	736.1729	0.135556	0.8927
NINETY	122.8277	35.80622	3.430347	0.0012
ADHERENTS	-2.633685	1.633244	-1.612548	0.1128
ADHERENTS^2	0.023276	0.019009	1.224444	0.2262
AFRAM	4.760926	9.649385	0.493392	0.6238

AFRAM^2	-0.025611	1.033214	-0.024787	0.9803
ASIAN	-13.78946	6.489340	-2.124939	0.0383
ASIAN^2	0.107356	0.504980	0.212595	0.8325
BACH	-0.787809	1.738781	-0.453081	0.6523
BACH^2	0.046088	0.031908	1.444410	0.1545
BACH*NINETY	-9.289659	3.059397	-3.036434	0.0037
(BACH*NINETY)^2	0.223729	0.061501	3.637827	0.0006
BELOWPOV	-5.903164	6.149709	-0.959909	0.3415
BELOWPOV^2	0.020289	0.181392	0.111851	0.9114
FAMILIES	19.42274	27.90544	0.696020	0.4895
FAMILIES^2	-0.122545	0.206415	-0.593683	0.5552
HIGH	-4.604730	9.508437	-0.484278	0.6302
HIGH^2	0.019088	0.062476	0.305528	0.7612
MARRIED	-7.702031	23.16856	-0.332435	0.7409
MARRIED^2	0.044445	0.205160	0.216638	0.8293
REALINCOME	-0.015356	0.012352	-1.243173	0.2193
REALINCOME^2	3.31E-07	2.72E-07	1.216330	0.2292
NATIVEAM	-1.466693	2.856563	-0.513447	0.6098
NATIVEAM^2	0.063031	0.146053	0.431561	0.6678

R-squared	0.576379	Mean dependent var	14.91181
Adjusted R-squared	0.392543	S.D. dependent var	26.18721
S.E. of regression	20.41018	Akaike info criterion	9.119808
Sum squared resid	22078.51	Schwarz criterion	9.850345
Log likelihood	-327.1126	F-statistic	3.135296
Durbin-Watson stat	2.127321	Prob(F-statistic)	0.000303

The abortion ratio:

Dependent Variable: RATIO
Method: Least Squares
Date: 12/10/02 Time: 10:06
Sample(adjusted): 1 78
Included observations: 77
Excluded observations: 1 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
NINETY	1255.520	349.4517	3.592829	0.0006
1-NINETY	1229.861	329.8367	3.728696	0.0004
ADHERENTS	-2.533081	0.896372	-2.825927	0.0063
AFRAM	26.32083	9.280938	2.836010	0.0061
ASIAN	-17.71086	9.387718	-1.886599	0.0638
BACH	-0.701407	1.980492	-0.354158	0.7244
BACH*NINETY	9.320086	2.096818	4.444872	0.0000
BELOWPOV	-3.461955	5.495324	-0.629982	0.5310
FAMILIES	-17.41941	7.219012	-2.412990	0.0187
HIGH	-1.440394	2.344118	-0.614472	0.5411
MARRIED	2.761717	5.920210	0.466490	0.6425
REALINCOME	0.013840	0.005276	2.623478	0.0109
REALINCOME*NINETY	-0.005960	0.005429	-1.097671	0.2765
NATIVEAM	2.487232	2.947132	0.843950	0.4019

R-squared	0.779350	Mean dependent var	249.9168
Adjusted R-squared	0.733820	S.D. dependent var	120.0720
S.E. of regression	61.94834	Akaike info criterion	11.25344

Sum squared resid	241768.6	Schwarz criterion	11.67959
Log likelihood	-419.2576	Durbin-Watson stat	1.661144

Heteroskedasticity test for the abortion ratio:

White Heteroskedasticity Test:

F-statistic	3.178298	Probability	0.000231
Obs*R-squared	46.89823	Probability	0.005040

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 12/10/02 Time: 11:05
Sample: 1 78
Included observations: 77
Excluded observations: 1

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-88959.77	155816.5	-0.570926	0.5706
NINETY	-20613.13	44896.98	-0.459121	0.6481
ADHERENTS	-740.9145	328.2554	-2.257128	0.0283
ADHERENTS^2	6.756484	3.806760	1.774865	0.0819
AFRAM	809.8301	1884.917	0.429637	0.6693
AFRAM^2	52.05668	200.9776	0.259017	0.7967
ASIAN	-3387.884	1277.414	-2.652142	0.0106
ASIAN^2	13.74981	98.14042	0.140103	0.8891
BACH	-227.7205	349.5477	-0.651472	0.5177
BACH^2	12.09221	6.471141	1.868636	0.0674
BACH*NINETY	-900.8635	651.6659	-1.382401	0.1729
(BACH*NINETY)^2	24.42763	12.71265	1.921521	0.0603
BELOWPOV	-216.9185	1218.571	-0.178011	0.8594
BELOWPOV^2	6.848158	35.71712	0.191733	0.8487
FAMILIES	11201.21	5558.958	2.014984	0.0492
FAMILIES^2	-84.15155	41.16542	-2.044229	0.0461
HIGH	-3069.940	1850.673	-1.658823	0.1033
HIGH^2	17.25279	12.16764	1.417924	0.1623
MARRIED	-2677.366	4498.836	-0.595124	0.5544
MARRIED^2	25.24814	39.82690	0.633947	0.5289
REALINCOME	-4.781849	3.455892	-1.383680	0.1725
REALINCOME^2	0.000112	7.44E-05	1.508651	0.1376
REALINCOME*NINETY	2.124746	3.896986	0.545228	0.5880
(REALINCOME*NINETY)^2	-3.35E-05	8.25E-05	-0.406277	0.6862
NATIVEAM	-836.2795	560.1794	-1.492878	0.1416
NATIVEAM^2	21.05169	28.93466	0.727559	0.4702

R-squared	0.609068	Mean dependent var	3139.852
Adjusted R-squared	0.417435	S.D. dependent var	5182.300
S.E. of regression	3955.441	Akaike info criterion	19.66692
Sum squared resid	7.98E+08	Schwarz criterion	20.45833
Log likelihood	-731.1763	F-statistic	3.178298
Durbin-Watson stat	2.012236	Prob(F-statistic)	0.000231

3. F-tests on statistically insignificant variables (at the 5% level)

The abortion rate:

Wald Test:
Equation: RATEREGRESSION

Test Statistic	Value	df	Probability
F-statistic	3.040478	(9, 64)	0.0043
Chi-square	27.36430	9	0.0012

Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
C(3)	-0.084240	0.061229
C(5)	-0.533672	0.637955
C(6)	-0.090255	0.132997
C(8)	-0.402001	0.308330
C(9)	-0.036655	0.466833
C(10)	-0.130025	0.157003
C(11)	-0.513638	0.382151
C(12)	0.000519	0.000361
C(13)	0.080666	0.192764

Restrictions are linear in coefficients.

The abortion ratio:

Wald Test:
Equation: RATIOREGRESSION

Test Statistic	Value	df	Probability
F-statistic	0.982587	(7, 63)	0.4520
Chi-square	6.878108	7	0.4417

Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
C(5)	-17.71086	9.387718
C(6)	-0.701407	1.980492
C(8)	-3.461955	5.495324
C(10)	-1.440394	2.344118
C(11)	2.761717	5.920210
C(13)	-0.005960	0.005429
C(14)	2.487232	2.947132

Restrictions are linear in coefficients.

4. Chow test for 1990/ 2000 data

The Chow test was performed on a "pooled" regression of observations for both years (with different intercept terms for 1990 and 2000) and two restricted regressions; one using only the 2000 data, and one using 1990 data only.

The following data are used to calculate the chow test (which assumes the null hypothesis that the difference between the coefficients over the two time periods is zero):

	residual sum of squares			number of observations		
	unrestricted	1990	2000	unrestricted	1990	2000
ratio	322349.7	149684.8	51032.43	77	39	38
rate	1350.04	829.5032	162.542	77	39	38

The chow test statistic is therefore

Ratio: ((322349.7-(149684.8+51032.43))/12)/(((149684.8+51032.43)/(77-24))

=2.6764 distributed f(12, 53)

Rate: ((1350.04-(829.5032+162.542)/12)/(((829.5032+162.542)/(77-24))

=1.59 distributed f(12, 53)

the f-distribution at the 5% level for 12, 53 degrees of freedom is approximately 1.96

Thus, we reject the null hypothesis that the slope coefficients are equal for the years 1990 and 2000 for the abortion ratio, and fail to reject the similar hypothesis for the abortion rate.