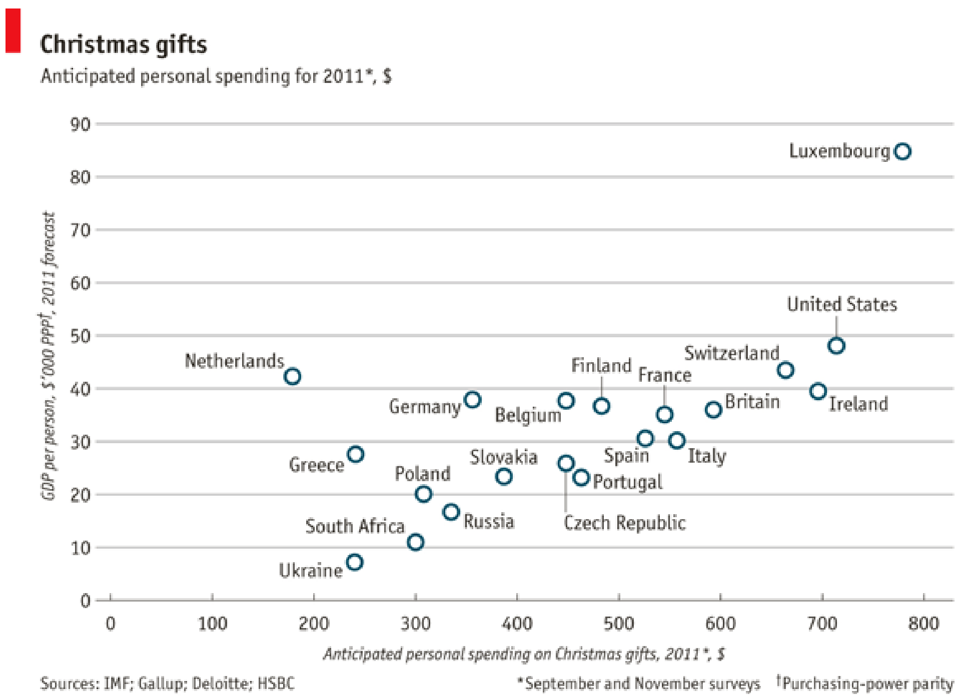
Final Project: Econ 312
Mark Jarrett and Stephanie Radoslovich
Jeff Parker

Introduction

This project is based off a data set from an *Economist* article. The article describes spending patterns of selected developed Christian nations based on GDP per capita. We wanted to examine other underlying factors that affect personal spending on Christmas. Below is the original graph.



Notable exceptions to the trend are the Netherlands, possibly due to its documented culturally ingrained frugality, and Luxembourg because of its incredibly high GDP per capita.

Data

The original data set included variables gifts and GDP per capita from 20 countries in the year 2011, as explained below.

Variable	Number of observations	Units
Gifts	20	Dollars per person
GDPPC	20	Thousands of dollars per person

The 20 countries included are the Netherlands, Greece, Ukraine, South Africa, Poland, Germany, Russia, Slovakia, Belgium, the Czech Republic, Portugal, Finland, Spain, France, Italy, Britain, Switzerland, Ireland, the United States, and Luxemburg. These countries represent most of the Christian world, excluding South America. We acquired six more variables from various sources to augment the data set. These variables are explained below,

Variable	Number of observations	Units	Source
Unemployment rate	20	Percentage unemployed	IMF
General government total expenditures	20	Percentage of GDP	IMF
Happiness net	15	Percentage of people who rated themselves as either "quite happy" or "very happy" minus the percentage of people who rated themselves as either "not very happy" or "not at all happy"	World Values Survey
Household expenditure on recreation and culture	16	Percentage of income spent on leisure	OECD
Sex ratio (M to F)	20	Ratio male to female	The World Factbook
Weekly church attendance rate	18	Percentage attending church	Nationmaster.com (questionable validity)

Below is Stata's summary of the dataset.

```
. summarize  gifts gdppc unemployment govexpend happiness rec sex church
```

Variable	Obs	Mean	Std. Dev.	Min	Max
gifts	20	460	172.5048	170	780
gdppc	20	33	16.2869	8	85
unemployment	20	.101253	.0539845	.03449	.24508
govexpend	20	.4486345	.0630079	.31773	.5651
happiness	15	.6193333	.3381561	-.04	.91
rec	16	.0485125	.0099691	.0279	.0645
sex	20	.9535	.0380132	.85	1
church	18	.3277778	.2166742	.02	.84

Results

```
. outreg, merge
```

	gifts	gifts	gifts	gifts
gdppc	23.915	6.402	5.971	7.187
	(1.76)	(1.70)	(1.91)	(3.92)**
unemployment	2,834.381	688.727	629.151	
	(1.72)	(0.69)	(0.69)	
govexpend	-755.469	-570.658	-573.403	
	(0.83)	(0.79)	(0.83)	
happiness	161.494			
	(0.44)			
rec	-8,458.656	-3,171.664	-3,202.562	
	(0.90)	(0.71)	(0.75)	
sex	-7,491.508	-643.183		
	(1.38)	(0.23)		
church	-32.125			
	(0.09)			
_cons	7,248.853	1,217.558	623.686	222.845
	(1.45)	(0.46)	(1.36)	(3.32)**
R2	0.59	0.45	0.44	0.46
N	12	16	16	20

* p<0.05; ** p<0.01

This outreg table demonstrates some of the faults in our original regressions. The adjusted R^2 values were -0.12, 0.17, 0.35, and 0.43 respectively. The initial negative adjusted R^2 value can be attributed to the missing data points in variables *happiness*, *church*, and *rec* almost halving our sample size. We dropped the most insignificant variables, which also increased our sample size, and reran the regression. The adjusted R^2 value became positive, but all variables were still insignificant. We dropped the lowest, *sex*, due to its insignificance and unusual coefficient, which gave us a huge bump to our adjusted R^2 in the third regression. For comparison, we ran a regression with only *gdppc* to see if our extra variables even warranted inclusion. Given gdppc's statistical significance when isolated, along with having the largest adjusted R^2 , we conclude that the extra variables had little to offer in terms of explanatory power. Next, we tried logging all variables. The results are shown below in a new outreg table.

```
. outreg, merge
```

	lgifts	lgifts	lgifts	lgifts	lgifts
lgdppc	**1.684**	**0.780**	**0.572**	**0.536**	**0.466**
	(2.34)	**(1.75)**	**(3.42)****	**(3.25)****	**(3.09)****
lunemp	**0.800**	**0.320**	**0.177**	**0.183**	
	(2.87)*	**(1.30)**	**(1.01)**	**(1.04)**	
lgov	**-1.706**	**-0.808**	**-0.599**		
	(2.02)	**(1.03)**	**(1.08)**		
lhap	**0.173**				
	(1.05)				
lrec	**-1.147**	**-0.384**			
	(1.53)	**(0.79)**			
lsex	**-18.977**	**-7.025**			
	(2.08)	**(0.97)**			
lchurch	**-0.184**				
	(1.11)				
_cons	**-3.522**	**2.048**	**4.062**	**4.684**	**4.481**
	(0.84)	**(1.13)**	**(5.12)****	**(8.51)****	**(8.69)****
R2	**0.74**	**0.42**	**0.43**	**0.39**	**0.35**
N	**12**	**16**	**20**	**20**	**20**

* $p<0.05$; ** $p<0.01$

The adjusted R^2 values are as follows: 0.29, 0.13, 0.32, 0.31, and 0.31. The first logged regression gave a huge R^2 and vastly increased the significance of the variables across the board, even with only 12 observations. The second regression was worse in all respects compared to the first and the third with no significant variables and lower adjusted and regular R^2 values. The third regression made our most logically significant variable *lgdppc* significant while still including other variables we find important with the largest adjusted R^2. It also includes *lunemp,* which the first regression found significant. The fourth and fifth regressions did not give any more information than the third. We thus chose the third logged regression as our model.

We tested for heteroskedasticity with hettest and found no evidence of its presence.

```
. hettest

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
         Ho: Constant variance
         Variables: fitted values of lgifts

         chi2(1)      =     0.01
         Prob > chi2  =   0.9083
```

Next, we ran a quantile regression to determine how the variables affect the output at the 25th, 50th, and 75th percentiles. This gives us a better understanding of which variables affect gifts on the lower, middle, and upper ends of the regression.

```
. outreg , merge
```

	lgifts	lgifts	lgifts
lgdppc	**0.583**	**0.532**	**0.617**
	(1.84)	**(2.99)****	**(8.45)****
lunemp	**0.047**	**0.100**	**0.040**
	(0.14)	**(0.54)**	**(0.52)**
lgov	**-0.742**	**0.052**	**-0.182**
	(0.71)	**(0.09)**	**(0.75)**
_cons	**3.520**	**4.627**	**4.111**
	(2.35)*	**(5.48)****	**(11.87)****
N	**20**	**20**	**20**

* p<0.05; ** p<0.01

These are the 25th, 50th, and 75th quantile regressions with pseudo R^2 values of 0.29, 0.41, and 0.50 respectively. This table shows that *lgdppc* is consistently effective, but increases with significance in the upper ranges. *Lunemp* is never statistically significant, but has over twice the effect on the middle range than on the upper and lower ends. *Lgov* has the greatest variance in its effect by far. It goes from being the largest absolute value coefficient on the lower end to a slightly positive coefficient in the middle, and it returns to negative on the lower end. However, none of these values are significant so we cannot be entirely sure of these effects.

Conclusion

The first thing to note about this project is its lack of observations. With only 20 countries, and some variables without the full 20 observations, it is difficult to create a compelling and statistically significant regression. We also believe that this analysis would fit much better with panel data instead of cross-sectional data. If we could observe the gift patterns of these countries over time, we could consider new variables and add weight to the ones we already have. The data we do have is from 2011, a time when most of these countries had suffering economies post-financial collapse, so this regression would not explain most other time periods. As this regression encompasses almost exclusively developed countries and requires that the countries celebrate Christian holidays, it would have little external validity.

In our chosen OLS regression, *lgdppc* has the strongest and most statistically significant effect on *lgifts*. We think this happens because when GDP per capita is higher, people buy more in general. *Lunemp* was also included in our chosen regression and has a positive affect on *lgifts*. This is unexpected, as we would think that more unemployed people would imply less gift buying. We also cannot forget that this coefficient was insignificant, and therefore we cannot say for certain that the positive effect is definite. The last variable included, *lgov*, has a negative but statistically insignificant effect on *lgifts*. Percent of GDP as government expenditure should have a negative effect on gift spending. A higher percentage implies consumption is a smaller part of GDP.

We omitted many of our variables due to statistical insignificance, but we included them originally to augment our regression. We suspected that higher happiness and general household recreation spending might have increased Christmas spending. The sex ratio's effect should have been negative as women are more prone to consumer spending. The church attendance variable would have been more interesting. More church attendance could have lead to a more purely religious view of Christmas, leading to less spending. However, more attendance could also have implied greater participation in Christmas, which might have lead to more spending. More countries might have made these variables more significant.

Our quantile regression provided much interesting information about how our variables affected gift spending. As GDP increases, people spend less of their money on necessities and more on less essential purchases, like gift spending. This fact explains why *lgdppc* has a stronger and more statistically significant effect on the upper end of the quantile regression. Unemployment has its strongest effect on the middle level of gift giving possibly because the countries at the top do not have much unemployment and the countries at the bottom do not spend as much anyway. Government spending is so strongly negative on the lower end because as the government spends more, especially in lower income countries, consumption might be decreased as a percentage of GDP. This might also imply higher taxes, which would also decrease consumer spending. On the upper end, people have more disposable income, so it would be less of a problem.

In conclusion, while this regression may not be very convincing due to the limitations on the data, it does give some indication that it could be redone with better data. It suggests that we might be able to more accurately predict gift spending in regards to our chosen variables, especially with panel data.