THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

DEPARTMENT OF ECONOMICS

AN EXPLANATION OF HOUSEHOLD FORMATION AND COMPOSITION IN DEVELOPED AND DEVELOPING COUNTRIES

ERICH MISKEWICZ
Spring 2012

A thesis
submitted in partial fulfillment
of the requirements
for baccalaureate degrees
in Economics and Finance
with honors in Economics

Reviewed and approved* by the following:

James R. Tybout
Professor of Economics
Thesis Supervisor and Honors Adviser

Bee Roberts
Professor of Economics and Asian Studies
Faculty Reader

* Signatures are on file in the Schreyer Honors College.

ABSTRACT

The landscape of households in the United States and much of Europe has been shifting toward smaller households, decreased marriages, and increased divorces over the past century. In the spirit of Gary Becker, this shift in household form can be attributed to the general decrease in relative price of labor saving home appliances. This paper, utilizing survey data from the Living Standards and Measurement Study, analyzes households in Albania to examine the relationship between the price of labor saving inputs into household production and household size. A model was developed that relates household size to consumption per member, mean education, the cost of energy, and the price of labor saving appliances such as washing machines, sewing machines, and dishwashers. There was found to be a significant positive relationship between the cost of washing machines and household size, suggesting that decreased gains from specialization may hold explanatory power even in developing countries.

TABLE OF CONTENTS

Section I - BACKGROUND

Introduction

The household, by some, is considered the most private and respectable sanctum in American life. However, the traditional household form has drastically changed in much of the developed world. In the United States the divorce rate had doubled from 1950 to 1990 and marriage rates have been declining depending on age (Greenwood and Guner, 2008). Female labor force participation had increased from 20% to 34.5% between 1900 and 1960 (Oppenheimer, 1967), and married women are now more likely to be the only working member in a household (US Census). In addition, unmarried couples have increased considerably in the United States, with the number of unmarried opposite sex couples doubling since 1999 (US Census). To understand these large changes to household composition, a greater understanding of the incentives and economic gains from household formation is necessary.

This paper will be examining one particularly important economic incentive that pertains to household formation: technological progress. As technology improves, home appliances such as washing machines or dishwashers become relatively less expensive. These household appliances lower the labor requirement for household production, inevitably lowering the incentives for household formation. The main relationship to be studied in this paper is the relationship between the prices of inputs into household production and household size, with the expectation that as the household production inputs lower in price, as will the average household size.

The importance of understanding the economic forces behind these changes witnessed in household composition is twofold. First, the issues of decreasing marriage rates and rising divorce rates are often discussed in the media and there seems to be a rising demand for the

return to the "traditional" family. However, it may be shown that these cultural shifts are actually responses to "naturally" occurring economic incentives. Second, household size has very important implications regarding poverty, as there is generally a negative association between household size and per capita income. Any research that can increase our understanding of the determinants of household size is very important regarding these areas of development economics.

The first section of this paper will begin with a basic setup of the framework of the household and household production. An understanding of this framework will allow for a better grasp of the incentives behind household formation and, more importantly, what factors play a role in shifting these incentives over time. Marriage will then be discussed, with an additional look at the rise of the importance of cohabitation in the United States and Europe. There will then be an examination of how household composition has been changing in the United States and Europe.

The second section of the paper will begin with a review the current empirical work on determinations of household size. A model of household size will then be developed, examining the causes of both fertility and marriage/cohabitation decisions that affect changes in household size. In order to better control for cultural differences, this model will be tested empirically using household level data from surveys by the Living Standards and Measurement Study conducted in Albania in 2005. The main purpose of this model is to find a possible relationship between household size and the price of inputs into household production.

The third section of the paper will summarize the results of the empirical work conducted in section II. It will discuss the strengths and weakness of this type of analysis, and offer suggestions for further research that could be designed and conducted to aid in finding the

solution to the question of whether the price of home appliances has any substantial impact on household size. The paper will conclude with an application of the results of this analysis to the overall trends seen in developed countries.

The household and household production

When discussing the concept of a household, it is best to start at its simplest form: a single-person household. Unlike traditional theory, in which a single person maximizes utility through the purchase of goods subject to a given budget constraint (Michael and Becker, 1973), the single-person household model "recognizes that each person allocates time as well as money income to different activities, receives income from time spent working in the marketplace, and receives utility from time spent eating, sleeping, watching television, gardening, and participating in many other activities" (Becker, 1991). This distinction is important because it allows for a separation of time spent in market work and time spent doing other activities.

Another important aspect of the single-person household model is household production. Household production is the concept that time and goods do not provide utility directly, but instead are considered as "inputs into the production of 'commodities', which directly provide utility" (Becker, 1991). These "commodities" can consist of "children, prestige and esteem, health, altruism, envy, and pleasures of the senses" (Becker, 1991) and are generally considered anything that bring about well-being or utility. The "commodities" cannot be purchased directly in the marketplace and are both produced and consumed by the household. For the purpose of this paper, it is important to note that the production of these "commodities" is the result of the input of some combination of both time and purchasable goods that is not necessarily constant over time.

Two centuries ago, many household in the United States produced a large fraction of what they consumed (Greenwood and Guner, 2008). On the contrast, modern households are able to replace much of the necessary time spent for household production with goods purchasable in the marketplace. Because of this, many households have a higher incentive to conduct market work, rather than spending time in domestic production. The recent shifts in household composition may be explained by shifts from the more rural households of the past to the more modern households of today.

Gains from household formation

The allocation of time to either market work or domestic production brings about large implications when discussing the division of labor within a multiple-person household. The separation of labor brings about the idea of "'production complementarities,' in which husband and wife specialize in the market and domestic spheres, respectively" (Stevenson and Woflers, 2007). These gains from specialization act as an incentive for individuals to enter a household. The division of market and household labor has historically been allocated via gender – men conduct the majority of the market work and women specialize in domestic production. However, according to Becker (1991), "even if a husband and wife are intrinsically identical, they gain from a division of labor between market and household activities." This implies that there need not be assumed biological differences between men and women for gains from specialization to be present; however, even a very small difference in ability based on gender, such as women being better at child rearing, would lead to a gendered division of labor in the household (Becker, 1991).

The gendered allocation of labor within a household is a very important concept, especially when examining the relationship between female labor force participation and

household formation rates. Sevilla-Sans (2007) found that "social norms toward the division of household labor may affect an individual decision to entering a household." In which case, less egalitarian norms may increase the perceived cost of men's domestic labor while lowering the cost of women's domestic labor. This opens up the framework for differences in household formation rates based on differences in social norms between cultures; however, the origin of these social precepts, as argued in this paper, could be of an economic nature.

In addition to gains from specialization in household production, Stevenson and Wolfers (2007) suggest that modern household formation is motivated by the existence of consumption complementarities and insurance benefits. Indeed, Nelson (1988) offers the theory and evidence for the existence of economies of scale in household consumption. In a study of five goods, Nelson discovered that there are "significant economies of scale with household size in consumption of all included goods." These economies of scale were found to be stronger for the consumption of shelter than for other consumption, such as that for transportation and clothing.

The major factors that provide incentive for household formation are gains from specialization in household production, economies of scale in household consumption, and other cultural or social incentives. When discussing the changing trends in household formation, it is important to examine what may cause shifts in the aforementioned incentives. The main focus of this paper will be on changes in the gains from specialization in household production, as this is an often cited cause of changes in household composition.

The basic principle to be examined is that technological progress will lower the gains from specialization (Greenwood and Guner, 2008). This is based on the fact that, as technology progresses, the prices of inputs into household production will become relatively lower. These inputs can include household appliances such as washing machines or dishwashers, prepared

food, as well as any purchasable services. A major assumption is that these purchasable inputs are used to replace the time that must be spent for household production. This reduction in the necessary time investment lowers the gains from specialization in household production, and ultimately lowers the economic incentives for household formation. This paper will attempt to quantify the relationship between the price of inputs into household production and household size.

Marriage and Cohabitation

Marriage has historically been the institution associated with family and the household. Becker (1991) has argued that because married women "have been specialized to childbearing and other domestic activities, they have demanded long-term "contracts" from their husbands to protect them against abandonment and other adversities." Marriage has, in the past, been treated as synonymous with household formation. Because cohabitation has become more prevalent in many modern societies (Bumpass, Sweet, and Cherlin, 1991), it is important to examine marriage as separate from household formation.

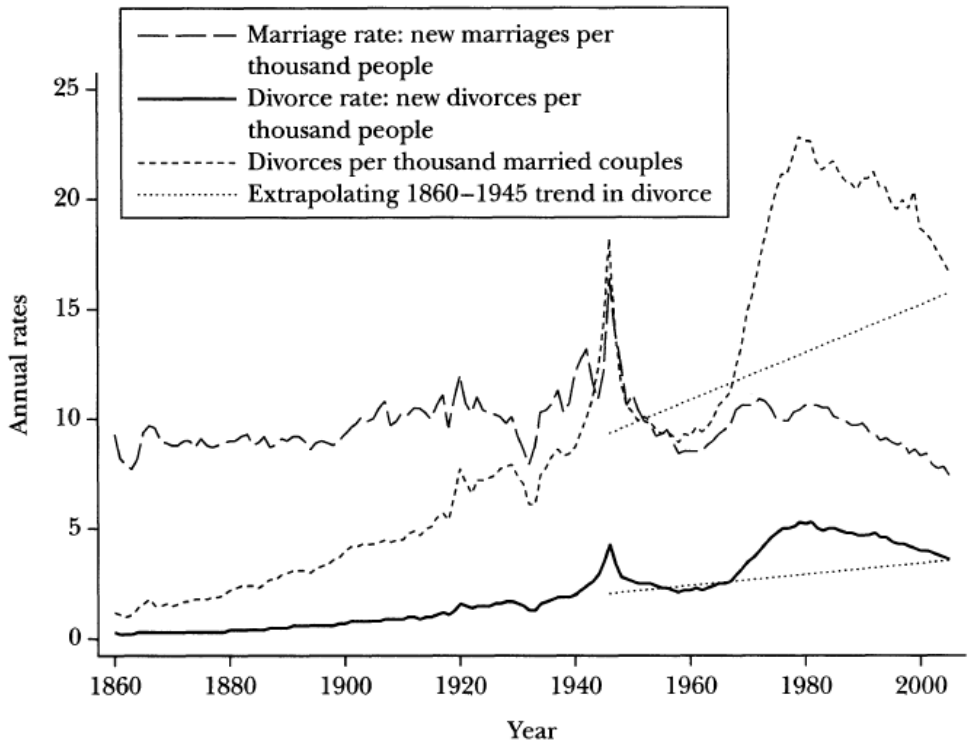
Marriage as an institution, independent of other economic gains, may have an impact on the incentives for household formation. Watson and Mclanahan (2011) studied the impact that identity and relative income have on marital status. They found that the decline in marriage in the United States since 1960 had become most pronounced at the bottom of the income distribution. They claimed that a male's "marriageability is contingent on steady employment on minimum level of earnings." This is derived from the importance of an individual identifying himself in the "married couple" category. They argued that this category brought with it a "prescription for a particular standard of living". This implies that marriage supplies a cultural incentive to be considered part of group that is held at a socially higher standard.

In the United States, a period of cohabitation is often seen as a temporary relationship stage that will result in eventual marriage (McGinnis, 2003; Cherlin 1990; Stevenson and Wolfers, 2007). Most cohabiting couples in the United States expect to eventually marry, despite that more than half of the couples that were cohabiting in January 1997 were no longer together five years later (Stevenson and Wolfers, 2007). In other countries, cohabitation is seen as a more permanent state. For example, in a 2003 survey, in Sweden only 31% agreed that "people who want children ought to get married," in comparison to 65.3% in the United States (Stevenson and Wolfers, 2007).

The shift toward cohabitation in many developed countries is an important change in household composition. If marriage functions as a contract to protect the household member with less economic power, a reduction in the traditional division of labor would decrease the necessity of such a contract and would ultimately have an effect on marriage rates. A shift toward increased cohabitation as a permanent stage is very consistent with the concept of decreased economic incentive for household formation and marriage.

Current Trends

The US has gone through large changes in household form. In 1950 82% of the female adult population was married, and in 2000 this number fell to 62% (Greenwood and Guner, 2008). The number of married females in the labor force has also increased from 23.7% in 1950 to 71% in 1990 (Greenwood and Guner, 2008). Stevenson and Wolfers (2007) discussed the changes in marriage and divorce rates in the United States over the past 150 years; the below graph depicts their findings.



Sources: Data for 1860–1919 are from Jacobson (1959); 1920–1998 from Carter et al. (2006); 1999–2005 from U.S. Census Bureau (2007).

[1]

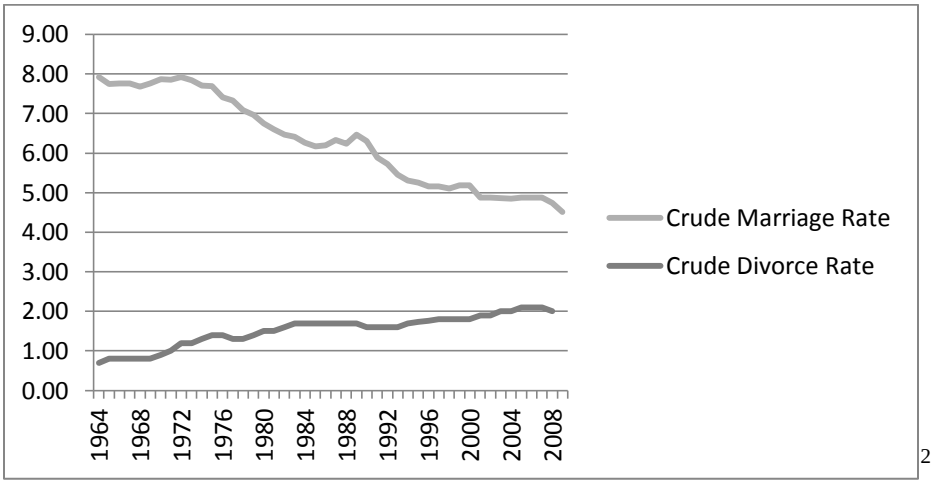
Marriage rates have remained relatively stable, though they have been decreasing since 1980. The Divorce rate has been following a general upward trend for the past 150 years, with a large spike between the mid 1960s and 1970s. This discrepancy is often cited to be explained by the liberalization of divorce laws in the 1970s; however, this holds little long-run explanation for increased divorce rates (Greenwood and Guner, 2007). Divorce rates have been decreasing since they peaked in 1981, perhaps returning to the original estimation based on an 1860 to 1945 trend in divorce (see above graph). Many attribute the large rise in divorce rates over this time to the rising economic opportunity for women, allowing them the freedom to escape bad marriages (Ruggles, 1997). Cherlin (1992) states that "almost every well-known scholar who has addressed [the rising divorce rates] in the twentieth century has cited the importance of the increase in the employment of women." The rise in divorce can also be explained in part by an

[1] Graph from Stevenson and Wolfers (2007)

"interdependence theory", which states that the benefits derived from specialization are lessened when both spouses engage in market work (Ruggles, 1997).

The trends in the United States do not differ from trends in other European countries. Since 1960 the number of marriages in the European Union has been falling steadily. From 1964 to 2009 the number of marriages in the countries in the European Union has fallen by 31% (Eurostat), while over the same time period the number of divorces in the EU had more than doubled (Eurostat).

Marriage and Divorce Rates in the EU



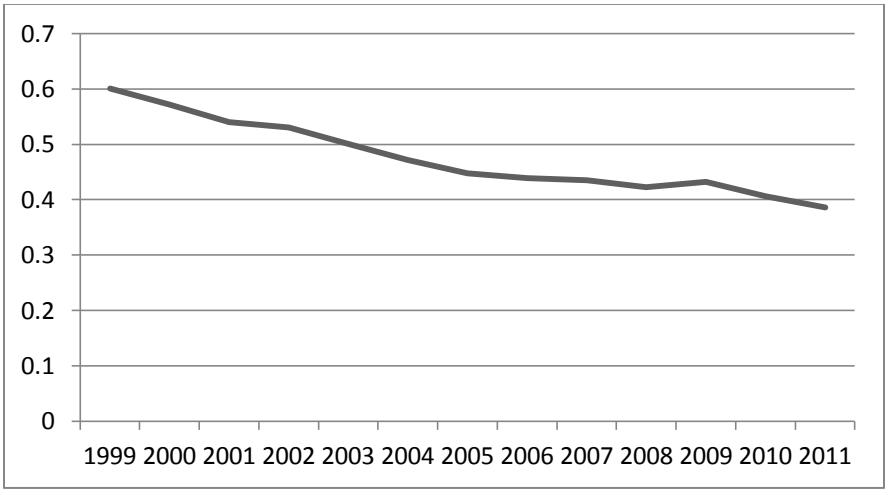
[2]

The attitude toward marriage also changed in many countries; in the UK and France 25.9% and 36.3% respectively agreed that marriage is an "out-dated institution" in 2003. This compared to around only 10% in the United States (Stevenson and Wolfers, 2007). However, in a study by Bumpass, Sweet, and Cherlin (1991), in the United States it was found that only 10% of those surveyed agreed that "It's better for a person to get married than to go through life being single." Though attitudes and social influences may be different between these countries, the general trends of increasing divorce rates and decreasing marriage rates are present.

[2] Data for graph from Eurostat

Since 1994 the average household size in the United States has fallen by 6.5%, from 2.67 to around 2.49 (Census). Over this same period, it is important to examine if there were any changes to prices of labor saving home appliances. To estimate these prices for the United States during this time period, the section in the Consumer Price Index labeled "Appliances" was examined. Over the time from 1999 to 2011, the price of appliances adjusted for inflation fell 36% (Consumer Price Index). In addition, the inflation adjusted price of "housekeeping supplies" has fallen by 17% since 1980 (Consumer Price Index). These statistics show that the prices of domestic labor saving appliances have been falling over much of the same period that household size has also been declining.

Relative Price of Appliances



[3]

Marriage, divorce, and household size indicators are impacted by many social and economic variables; however, the overall trends of shifting household form are existent across cultural and geographical barriers. To better understand the impact that changes in the gains from specialization in household production have on these variables, it will be necessary to

[3] Data for graph from Consumer Price Index

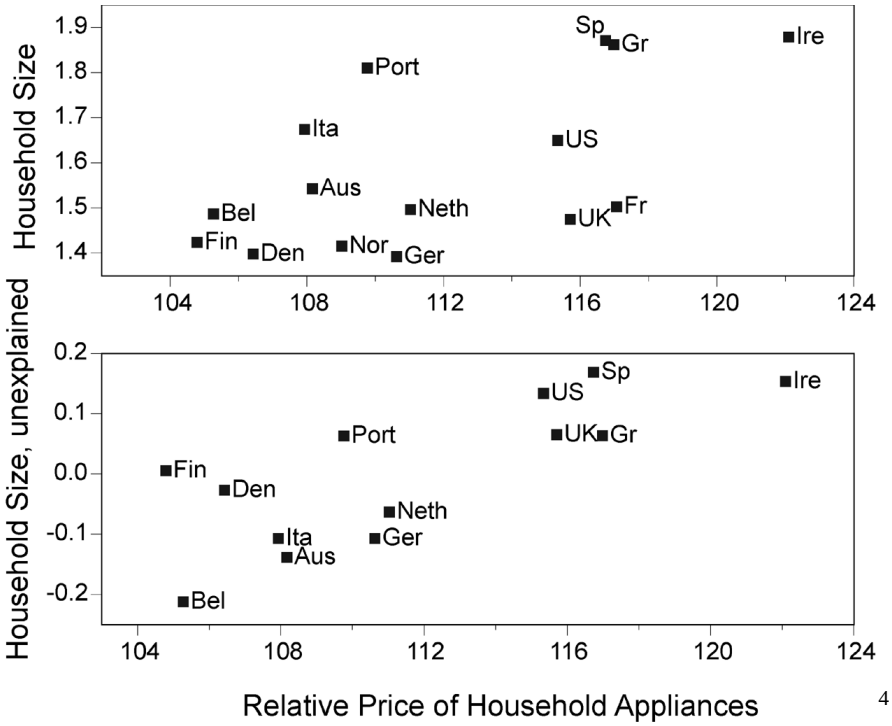
examine the changes on a microeconomic level to better control for the vast cultural differences between these countries.

Section II – EMPIRICAL ANALYSIS

Summary of Relevant Empirical Work

There has been limited empirical work done on the direct relationship between relative price of household appliances and household size. However, below is a summary of work that has been done to explain shifts in household size, as well as work that is relevant to general changes in household composition in the United States over the past century.

Greenwood and Guner (2008) examined cross-country relationships between the relative price of household inputs into household production and household size. Using a sample of 10 western countries, they calculated a linear relationship between relative price of appliances and household size. In the analysis they included other control variables such as "GDP per capita, the gender gap, the extent of urbanization, the amount of product market regulation, and the religiosity of a population." In their results they found a statistically significant positive relationship between the relative price of appliances and household size. In addition, they found the "price effect" to be quantitatively powerful, explaining about 100% of the observed difference in household size between Finland and Ireland, which had the highest and lowest relative price levels. In comparison to the other variables, the authors came to the conclusion that the relative price of household appliances "appear(s) to have a strong impact on household size." The below chart exhibits the relationship between household size and relative prices for 15 different countries.



[4]

The aforementioned results must be interpreted cautiously, as recommended by Greenwood and Guner, due to the small sample size of the analysis. However, it is important to note also that an empirical analysis on such a large scale is vulnerable to problems of endogeneity. Due to large immeasurable cultural differences between the countries in the study, it may be impossible to control for all of the factors that have an impact on household size, and therefore, the results of Greenwood and Guner may also suffer from omitted variable bias. However, based on the relative strength of their results, it seems to be strong evidence for the existence of a positive relationship between the relative price of household appliances and household size.

What is to be determined is whether or not this relationship remains true for households in poorer countries. There may be certain differences between developed and developing countries that diminish the relationship explained by Greenwood and Guner. In developing

[4] Graph from Greenwood and Guner (2008). "Household Size, unexplained" is the deviation of household size predicted by the other controls from the actual household size measured.

countries, for example, the prospect of remaining single is much less viable, and measures of household size may be much more dependent on fertility rates than marriage or cohabitation rates. For example, from a survey of households in Albania, only roughly 13% of the households did not have a member who was considered the "spouse/partner" of the household head (LSMS). Additionally, according to Easterlin, Pollack, and Wachter (1980), there exists a "general absence, rather than presence, of deliberate fertility control in less developed countries." This absence of fertility control supports the idea that household size may be more a function of fertility in developing countries. However, the degree of this difference between developed and developing countries is uncertain, and a causal relationship between relative price of household appliances and household size may still be present to some extent.

Ruggles (1997) examined the effects of different variables on marriage stability in the United States between 1880 and 1990. The author found that while age had a relatively weak effect on marriage instability, the magnitude of the coefficients on labor-market variables was very significant. Indeed, there was found to be a strong positive correlation between both male and female labor force participation and the probability of being divorced or separated. His findings also contribute to understanding why marriage instability has generally been higher among the black community. This is based on the fact that "blacks have significantly higher female participation and lower male economic opportunities" relative to whites. These findings support the overall theme of the observed changes in household form in the United States over the previous century.

With regards to marriage, McGinnis (2003) developed a model for the probability of marriage in cohabiting couples based on "costs and benefits of marriage, marriage expectations, [and] marriage intentions." The model was analyzed using data from a survey that included 682

cohabiting couples. The author found that cohabiters perceive fewer costs to marrying than couples not living together, and employed persons see "significantly fewer costs to marriage than those not working for pay." The study also found that cohabiters perceive less benefits to marriage than noncohabiters, while at the same time having higher intent and expectation for marriage. Ultimately, the author found that probability of marriage could be "directly predicted by marriage intentions, marriage expectations, and the perceived costs of marrying." As argued in this paper, these intentions and expectations may be shaped by economic incentives.

Musgrove (1980) analyzed the determinants of household size and composition in Latin America in 1967-69. Using information from a household budget survey, it was found that "family composition alone is at least as strongly related to poverty as variables involving employment and dependency rates" and that "poor tend to be found in larger families." An important implication of this study is that household size is negatively correlated with income, or consumption, per person. Musgrove continues to discuss the reasons of this correlation by explaining the "life-cycle" effect. This effect states that as children are added to a family, consumption per person will inevitably decrease, causing a "period of relative poverty at one stage of the life cycle." While Musgrove's focus was primarily on poverty, his conclusions are important in understanding the relationship between household size and income/consumption, especially when examining poor families.

Data description

As stated previously, to examine the issue at its core, data on the microeconomic level has been examined to better control for cultural and other dissimilarities between populations. The Living Standards and Measurement Study (LSMS) available through the World Bank is a

viable option for an examination on this level. Satisfactory data is obtainable for households in Albania, for the year 2005.

The LSMS provides data collected through household questionnaires on daily consumption, community, dwelling descriptions, and pricing. The study also provides information on the location of each household, which includes district, municipality, as well as longitude and latitude. In the surveys, household members are considered those who have been away from the household for less than six months. The surveys identify a household "head" and all other individual household members. Though the assumption of the existence of a "head" of household may not be applicable in some developed countries, it remains much more relevant in developing countries and is unlikely to cause any issues with data collection or bias. Additionally, it highlights cultural differences of how household composition is viewed between developed and developing countries.

1. Variables

To measure the relationship between household size and the relative price of household inputs to production, it is important to capture all relative variables that may influence household size. For the purpose of this study, household size will be measured as the total number of members residing in the household, regardless of age. A representative of the relative price of household inputs to production will be a more difficult variable to produce. The household questionnaire includes information of expected resale value of various owned appliances. The appliances that most resemble inputs into household production in this case are Washing Machines, Sewing Machines, and Dishwashers. This information, in conjunction with information of household location, will allow for the calculation of an average cost for each relevant household appliance in each of the 36 districts included in the survey. Utilizing this

average district cost will help diminish problems regarding the perceived appliance value that may differ between individual households within the same district.

Based on the fact that each of these appliances requires energy to operate, it is advisable to include the cost of energy that each household faces in the analysis. It is clear that higher energy costs would raise the relative cost of any appliance that uses energy. It can also be expected that higher energy costs would lead to larger gains from economies of scale in household consumption, due to relatively higher benefits from dividing the cost of electricity. The LSMS surveys provide information on the cost as well as the availability and reliability of energy to each household. A household that has a lower availability to or a less reliable energy source will also be less likely to own these household appliances, as the perceived benefits will be lower. Therefore, in households that face these factors, the labor saving effect of the appliances would be relatively lower. As with the cost of household appliances, the cost of energy faced by each individual household will also be calculated by using district averages across households to regulate household estimations.

2. Controls

It will be important to include other measurements provided in the surveys that are relevant to the size of the household. According to Lanjouw and Ravallion (1994) there is "considerable evidence of a strong negative correlation between household size and consumption (or income) per person in developing countries," and this relationship is backed up by other sources (Musgrove, 1980; Cutright 1971). There is also further evidence of the possible relationship between education and fertility (Martin and Juarez, 1995; Rindfuss and Bumpass, 1978), particularly through implications of greater understanding of birth control and higher aspirations and career goals for women. It is also expected that any assistance received from

outside sources may affect decisions made about household formation or fertility. For these reasons, the other important variables that will be included as controls for the analysis are income, education level, and social assistance.

With regards to education level, the questionnaire includes information on literacy, number of years of schooling, degrees attained, private tutoring, and the cost of schooling. In order to calculate the education level for an individual, the "highest grade level reached" will be used. For the education level of a household, the mean adult grade level reached is utilized, with an adult being any individual over the age of 18. However, it is important to note that the education level of the women in a household may be more relevant in the prediction of fertility and, therefore, household size.

Household income will also be included. The surveys provide information on occupational wage, non-farm business profits, and other income. The total household income will be used for the analysis, which will be calculated as the sum of all individual incomes within the household. A measurement of social assistance will also be included, as any additional assistance would certainly have an impact on the gains to household formation. However, due to the fact that the amount of social assistance a household receives would most likely be negatively correlated with the household's level of income, social assistance may be considered as an additional income for the household, and added into the calculation for total household income. Because a larger household would be expected to have more income, the regression will use a per capita income to avoid this dual causality.

In Albania a large proportion of the households are agricultural, and therefore report zero "non-farm" income in the survey. Indeed, only about 35% of the households surveyed report any income at all. Conversely, almost 50% of survey respondents report owning or renting some

amount of land for agricultural use or animal husbandry. Being that agriculture plays a large role in the generation of income and food for many of the surveyed households, income generated through agricultural means will be included in the summation of household income. This will be estimated using the "net amount… obtained from the sale of [agricultural products] during the last 12 months". For agricultural byproducts, this measure is an estimate after subtracting the cost of inputs.

One issue with this method is it only includes the profit that households have earned from their land. According to the surveys, about 99% of households owning land earned some profits from the sale of animals or other harvests. The remaining households produced goods only for their own consumption, and, as this is not considered income, would not be included in the analysis. In addition to these households with zero income, many of the households that did profit from their land also produced for their own consumption, and this income, while important, would not be included. For this reason, the market value of all goods produced (even those for consumption) must be estimated and included as income for each household. This can be achieved by calculating an average price for each individual good using the data from the households that sold the good. Then, using these average prices, an estimate for the market value of every good that a household had harvested throughout the year can be obtained. This method was used to sum the income from annual harvests, tree harvests, and animal and agricultural byproducts. The usage of an average price may not be ideal because different regions of the country most likely face different prices; however, the number of observations by district/region is limited, so the overall average is the best course of action. Also included is the income received from sale of animals; this was reported as a nominal amount in all cases, so no use of averages was necessary.

Possible Measurement Issues

There are various possible measurement problems with the data being utilized for the analysis. Due to the fact that the data is collected through household questionnaires and is completed in two separate visits, refusal and absence may always impact the final results. There is also an issue in that much of the data represents the household's perceived value of certain items rather than the actual value that the household may face. The use of average prices in different districts may reduce the bias, though the lack of accuracy in the reported figures still may impact the overall results. There are also questions in which respondents must estimate incomes received over long time periods (up to 12 months). These questions are subject to even further guesswork and misreporting of facts. There may also be bias in the way individuals may overestimate their reported income, or any other nominal factors. Finally, though the study examines households on a microeconomic scale within the same country, there still may be cultural and social differences between the households not captured in the survey.

Descriptive Charts

Variable by district	Obs	Mean	Std. Dev.	Min	Max
Sewing machine price	35	109832.2	45174.06	43250	275000
Dishwasher price	10	196666.7	129843.2	20000	400000
Washing machine price	35	233913.7	77604.53	60964.29	425000
Cost of energy	36	81.96486	19.56866	49.01587	130.1334

The above chart represents the summary statistics for the average price of sewing machines, dishwashers, and washing machines in the 36 districts available for the study. Note that information on the price of dishwashers was limited to only 10 districts because many fewer households owned dishwashers than the other two appliances. This may be because the marginal benefit of owning a dishwasher is lower and not worth the relatively high cost of the appliance. For this reason, when including the dishwasher price in the relative cost of household inputs to

production, the sample size of the analysis will be limited to only the districts where data is sufficiently available. Many of the districts also have very few observations for the prices of washing machines and sewing machines. It is important to note the extreme difference between the districts with the minimum and maximum prices of the appliances. This may be evidence of a very large gap in the prices faced by those living in different regions of Albania. Also included is the cost of energy for each of the 36 districts. This value is cost in Leks per KWH and was calculated by dividing the estimated amount of energy used during a billing period by the total amount paid. The survey supplied ranges of energy used, and for the calculations the difference was split: e.g. 101-200 KWH = 150 KWH, etc.

Variable	Obs	Mean	Std. Dev.	Min	Max
Household size	3840	4.505729	1.805311	1	16
Mean schooling	3840	5.41392	1.783526	0	8
Total non-farm income	3840	1,562,530	1.23e+07	0	4.80e+08
Social assistance	3840	48,008.54	66041.33	0	1600000
Total Agricultural Income	3840	1691552	3924919	0	1.21e+08
Average Income	3840	766920.4	3402680	0	1.54e+08

The above chart shows the summary statistics for the other important variables being considered. The mean household size of the sample is 4.5, which is much greater than that of the United States and many European countries (Census; Eurostat). Mean schooling is represented by the average highest grade reached by individuals within a household. In this case, on average, the highest grade reached is approximately 5.5. The income and social assistance measures are the total amounts for the household. The values represent a year's income and all social assistance received in the 12 months prior to the survey. These values are in the Albanian currency, the Lek, which, during the time of this analysis, was trading at 105.75 Lek per USD. Average Income is a measure of income (non-farm and agricultural) per member of the household.

The Model and Regressions

Based on the above discussion, household size will be treated as a dependent variable and will be calculated as having a log-log relationship with relative price of household inputs to production and an array of various other controls. A log relationship will be utilized because it is likely that items such as income and price of household inputs effect household size by percent changes rather than nominal changes. The general model is as follows:

$$\text{HOUSEHOLD SIZE} = \alpha + \beta 1\ \text{PRICE OF INPUTS} + \beta 2\ \text{CONTROLS} + \varepsilon$$

PRICE OF INPUTS is a variable that represents the average relative cost of labor saving appliances faced by a household and CONTROLS will include various combinations of the other variables mentioned earlier in the paper.

It is expected that the coefficient on the price of inputs will be positive, i.e. as labor saving appliances become relatively less expensive, household size will tend to decrease. The coefficient on education level is expected to be negative, based on the prediction that higher educated adults will be less likely to have children due to relatively lower perceived benefits. It is also expected that income will be negatively correlated with household size because households with higher income will also experience relatively lower gains from having children. The cost of energy is expected to be directly related to household size for two reasons. First, as energy becomes more expensive, returns to scale will increase. Second, the higher the cost of energy, inputs into household production become relatively more expensive, which implies household size will tend to increase.

1. Model 1

The first model was regressed as log(householdsize) = $\alpha + \beta_1$*log(average income) + β_2*log(mean schooling) + β_3*log(energycost) + β_4*log(washing machine price) + β_5*log(sewing machine price) + β_6*log(dishwasher price). Note the results below:

	Coef	Std. Err.	t	P>|t|
log(average income)	0.010852	0.001729	6.28	0.000
log(mean schooling)	0.23033	0.033793	6.82	0.000
log(energy cost)	-0.14411	0.077308	-1.86	0.062
log(washing machine price)	0.188894	0.049683	3.8	0.000
log(sewing machine price)	-0.09168	0.044381	-2.07	0.039
log(dishwasher price)	-0.03655	0.026001	-1.41	0.16
constant	0.800766	0.818395	0.98	0.328
observations	1849			
Adj. R-squared	0.0664			

As expected the coefficient on the price of washing machines was positive and significant. However, many of the results were unexpected. The coefficient on average income was positive. This result is the opposite of the expectation and contradicts what traditional theory would suggest. Similarly, the coefficient on mean schooling was also positive and significant, suggesting more educated adults tend to live in larger households. Energy cost had a negative coefficient and was found significant on a 90% level. This suggests that economies of scale in consumption of energy may not offer enough incentive for household formation or that the cost of energy has a negative effect on fertility decisions. Contrary to the price of washing machines, the coefficients on the cost of sewing machines and dishwashers were found to be negative and less significant. Also note the low Adj. R-squared value of .0664, suggesting this model has little predicting power.

The model as it stands has various issues. The first issue is that including the price of a dishwasher in the analysis excludes many of the observations in the districts where the price was unavailable. Also, the price for some districts was estimated based on very few observations and may be inaccurate. For these reasons, it may be advisable to exclude the price of dishwashers from future regressions. Additionally, there could be the existence of multicollinearity between

the prices of the appliances and between income and education, but this does not lower the overall predictive power of the model.

Another serious issue with the model pertains to the data. Even after including the market value of all crops produced in total income for each household, roughly 25% of the households still have zero income. In this case, a household with zero income means that the household has zero consumption, which is clearly impossible. This issue could arise for a number of reasons. The data collection could be flawed and certain food production may be miscalculated or missed completely; there may also be flaws in the method used to calculate the per household consumption. There are two ways that this issue can be reduced. First, with social assistance factored into income, the percentage of households with zero income lowers. Second, the households with zero income can be excluded from the regression. While this is not the best method to solve the issue, it may still provide further insight into the relationship of the variables.

2. Model 2

The second model includes social assistance as household income and was regressed as log(householdsize) = α + β1*log(average income and social assistance) + β2*log(mean schooling) + β3*log(energy cost) + β4*log(washing machine price) + β5*log(sewing machine price). Note the results below:

	Coef	Std. Err.	t	P>|t|
log(average income and sa)	0.002795	0.0016	1.75	0.081
log(mean schooling)	0.279012	0.023041	12.11	0.000
log(energy cost)	-0.08695	0.03732	-2.33	0.02
log(washing machine price)	0.1381	0.030005	4.6	0.000
log(sewing machine price)	0.006357	0.02586	0.25	0.806
constant	-0.43874	0.466501	-0.94	0.347
observations	3732			
Adj. R-squared	0.0579			

When social assistance is included, the relationship between income and household size is positive and no longer statistically significant on the 95% level. The coefficients on education and the price of washing machines, however, are still positive and significant. The sewing machine price no longer has any significant relationship with household size. These results are surprising because, based on traditional expectations, income should be negatively correlated with household size. It may be the case that the receivers of social assistance are somehow determined independently of household size, and this dampens the observed effects of income. Additionally, social assistance may be considered a short term income, and therefore has little effect on long term decisions (such as fertility and household formation). However, even with social assistance factored in, there still remain a number of households with zero income. For this reason, the next model will exclude these observations.

3. Model 3

The following regression includes only the observations which report positive income (average income>0), and excludes the price of dishwashers from the explanatory variables. The model was regressed as log(householdsize) = $\alpha + \beta_1$*log(average income) + β_2*log(mean schooling) + β_3*log(energycost) + β_4*log(washing machine price) + β_5*log(sewing machine price) if average income > 0. Note the results below:

	Coef	Std. Err.	t	P>|t|
log(average income)	-0.03903	0.006062	-6.44	0
log(mean schooling)	0.295855	0.025885	11.43	0
log(energy cost)	-0.08731	0.04065	-2.15	0.032
log(washing machine price)	0.116775	0.030584	3.82	0
log(sewing machine price)	-0.0057	0.026594	-0.21	0.83
constant	0.531551	0.500356	1.06	0.288
observations	2553			
Adj. R-squared	0.0824			

The first result that stands out is the coefficient on average income is now negative and significant. While this result is more in line with traditional expectations, it is clear the exclusion of the observations reporting zero income had a strong impact on the results. However, it is safe to say that, of the households with positive income, households with higher average consumption tend to be smaller. The coefficient on mean schooling remains positive and significant. Also interesting, there is an observed negative relationship between energy cost and household size that is significant on the 95% level. A negative relationship is the opposite of what was expected due to economies of scale in household consumption, so this regression suggests that there may be some other way that energy cost could be influencing fertility or marriage/cohabitation decisions. There could also be some correlation between energy cost and other variables, i.e. households in districts with higher energy costs tend to have higher incomes or be faced with higher priced washing machines.

The coefficient on the price of washing machines is still positive and significant, while the price of sewing machines remains unrelated to household size. It is also important to note the adjusted r-squared of this regression, while still low, is higher than the first model, suggesting that changes between the two regressions made an improvement in the predictive power of the model.

The model, while suffering from the obvious issues of selection bias (ignoring households with zero income), is most in line with expectations. As stated previously, there should most likely be a negative correlation between average consumption and household size. Because model 3 expresses this negative relationship, it can be argued that of the two methods of dealing with the households with zero income, excluding the observations was the most effective method. However, the exclusion of observations with zero income may ultimately render the

relationships measured inaccurate. Despite this, the information portrayed has some degree of accuracy for households with positive incomes. Also, when taken in conjunction with Model 1 and Model 2, some important conclusions are able to be drawn from the results.

4. Regression Chart

The below chart contains the results from various regressions including only the observations in which household income is greater than zero. Log(household size) is the dependent variable in all of the regressions.

Regression	1	2	3	4	5	6	7	8	9	10	11
log(average income)	-.0390*	-.0363*	-.0309*	-.0394*	-.0341*	-.0351*	-.0357*	-.0372*			
log(mean schooling)	.2959*							.3028*			.3245*
mean schooling		.0711*	.0634*		.0728*	.0714*	.0719*		.0729*	.0741*	
log(energy cost)	-.0873**	-0.0583	-.1643**								
log(washing machine price)	.1168*	.1116*	.1496*			.1178*	.1230*	.1341*	.1139*		
log(sewing machine price)	-0.0057	-0.0151	-.1296*				-0.0102				
log(dishwasher price)			-.0548***								
constant	0.5316	0.6275	2.6086*	1.9996*	1.5210*	0.1018	0.1591	-.1689	-0.3217	1.0642*	.9445*
Adj. R-squared	0.0824	0.111	0.0917	0.0137	0.1024	0.1093	0.1107	0.0805	0.0985	0.0922	0.0587

*significant on a 99% level	**significant on a 95% level	***significant on a 90% level

In examination of the above chart, a few conclusions are able to be drawn. The coefficient on average income is negative and significant across the board. It is also clear that mean schooling tends to have more explanatory power than the log of mean schooling. Energy cost and the price of sewing machines may have some explanatory power of household size, as the adjusted R squared tends to be higher when they are included (regression 2 having the highest adjusted R squared). Also very important, the coefficient on the price of washing machines remains positive and significant across all regressions. The consistency of the results suggests that the correlations observed between these variables and household size are significant and offer at least limited explanatory power of the dependent variable.

It is also important to observe the size of the coefficient on the log of the price of washing machines. Excluding the log of mean schooling, it is larger than most of the other variables.

This implies that the price of washing machines is a relatively important variable in the determination of household size compared to average income.

5. Other Models

The results of other regressions are reported in the appendix. Regression A and B are used to measure the relationship using household size as a dependent variable, rather than log(household size). The results of these regressions are similar, though the relationships are not as strong. These results suggest that the log-log relationship is the best at explaining the relationship between the variables.

Regression C takes a different approach, and analyzes the relationship between the variables and only the number of adults in the household. The idea is that not counting individuals under the age of 18 may control for fertility differences, and only focus on household formation through marriage and cohabitation. The results, however, are unsatisfactory, and none of the coefficients on the independent variables were significant, except the coefficient on mean schooling, and the model as a whole had very little explanatory power.

Finally, Regression D includes only market income for a household; this excludes the value of any crops grown only for consumption and is also more accurately measured because all market income was reported nominally. The results, however, were generally similar to the results found when all consumption goods are factored into income.

Section III – CONCLUSION

Model Conclusions

According to the economic theory of household formation, it is to be expected that a decrease in relative price of inputs into household production will result in lower household size. The findings in this paper seem to confirm a positive correlation between the price of washing machines in 31 districts in Albania and the number of members in a household based on the vast consistency of results across various models. No relationship was confirmed, however, between the prices of other inputs into household production such as sewing machines and dishwashers, except for maybe a small negative correlation in some cases. This relationship was found after controlling for other factors such as average consumption, education, and energy cost. Greenwood and Guner (2008) had shown empirically that this relationship is likely to exist in a small sample of developed countries. This paper has shown that even when examining individual households within a developing country, there is some evidence for a causal relationship between the price of inputs into household production and household size.

Why would the price of washing machines be relatively more important than the price of sewing machines and dishwashers? For one, sewing machines, and dishwashers to a lesser extent, make up a much smaller proportion of a household's budget, with sewing machines costing roughly half that of washing machines. This lower relative price may lessen the effect that the input prices have on household size. The relationship may also be much weaker for sewing machines and dishwashers because of the relative labor saving benefit versus the cost of the items. While a washing machine costs more, the amount of labor that it saves is many times that of a sewing machine or dishwasher. This is clear when you compare the tasks of hand washing clothing to hand washing dishes, the former being highly labor intensive. When

considering a ratio of labor saved over cost, the price of a product with a higher ratio would be relatively more important in the determination of household size than the price of a product with a lower ratio. With this in mind, it is understandable that the "washing machine" effect would be much more visible than the "sewing machine" or "dishwasher" effect.

The results, however, must be interpreted cautiously. Though there appears to be a positive correlation, it may be difficult to draw the conclusion that the price of washing machines truly causes household size. The statistical methods used to determine this relationship means the conclusions may lack internal validity. The data utilization method may also lead to flawed results through the assumption that every household in a district is faced with the same price of washing machines, sewing machines, and dishwashers. There could indeed be some other factor associated with the district that affects both the price of washing machines and the size of households within the district. Additionally, controlling for all factors that play a role in fertility and household formation decisions may require information not provided in the surveys utilized in this analysis.

Future Research

Utilizing the current survey data, the research could be expanded to include additional countries and years to see if the relationship exists on a larger scale and show that Albania is not just an exception to the rule. It could also be possible to utilize different empirical methods. Rodrik (2008) contains a relevant discussion on the pros and cons of randomized control experiments versus larger scale empirical evaluations in what he refers to as "the new development economics". To examine the causal relationship, if there is one, between the price of inputs into household production and household size, a legitimate randomized control experiment could be performed; however, the difficulty and cost of performing such an

experiment versus the relative gains from the experiment make this idea unlikely. Any results from such an experiment would also be difficult to generalize, as it would be hard to prove that any given sample is representative of the population. For the above reasons, randomized control experiments may not be the correct route for further research.

The results could also be aided by the development of a survey that asks more specifically about inputs into household production. This could include questions about specific appliances such as "if you bought a washing machine today, how much would pay?" The survey could also include helpful information on the utility of these appliances with questions such as, "How much would you be *willing* to pay for a washing machine?", "How long would it take you to hand-wash your clothing?", etc. A survey such as this one could provide useful information on not just the cost of appliances, but also how helpful such appliances are to individual households. This would help alleviate some of the differences between households pertaining to use value and exchange value.

Summary of the Relevance of the Results

The changing landscape of households in the United States and much of the rest of the developed world are important to our cultural and economic future. Aspects of this new cultural landscape, such as below replacement fertility and changing attitudes toward marriage and cohabitation, can be seen as the results of decreased gains from specialization within a household. Even in a developing country, this effect may be witnessed. Though weaker, these economic forces may play a role in driving the cultural attitudes and decisions of the people. Indeed, it may be possible to show that economic changes that incentivize these shifts may precede any cultural changes, and, therefore, what are most often referred to as "cultural" changes could be seen as responses to changes in incentives.

Appendix

Regression A

```
      Source |       SS       df       MS              Number of obs =    2553
-------------+------------------------------           F(  5,  2547) =   51.33
       Model |  738.346007     5  147.669201           Prob > F      =  0.0000
    Residual |  7327.27561  2547  2.87682592           R-squared     =  0.0915
-------------+------------------------------           Adj R-squared =  0.0898
       Total |  8065.62162  2552  3.16051004           Root MSE      =  1.6961

------------------------------------------------------------------------------
householdsize |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
--------------+---------------------------------------------------------------
  logavincome |  -.1687058   .0261876    -6.44   0.000     -.220057   -.1173546
logmeanschool |    1.31961   .1118191    11.80   0.000     1.100344    1.538875
logenergycost |  -.2854943   .1756048    -1.63   0.104    -.6298369    .0588484
   logwmprice |   .6547869   .1321194     4.96   0.000     .3957145    .9138593
   logsmprice |  -.0154233    .114884    -0.13   0.893    -.2406989    .2098523
        _cons |  -1.775193   2.161488    -0.82   0.412    -6.013645    2.463259
------------------------------------------------------------------------------
```

Regression B

```
      Source |       SS       df       MS              Number of obs =    3792
-------------+------------------------------           F(  5,  3786) =   88.61
       Model |  1303.67488     5  260.734976           Prob > F      =  0.0000
    Residual |  11140.1856  3786  2.94246847           R-squared     =  0.1048
-------------+------------------------------           Adj R-squared =  0.1036
       Total |  12443.8605  3791  3.28247441           Root MSE      =  1.7154

------------------------------------------------------------------------------
householdsize |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
--------------+---------------------------------------------------------------
  logavincome |   .0429783   .0046321     9.28   0.000     .0338966      .05206
   meanschool |   .2394654   .0159477    15.02   0.000     .2081985    .2707323
logenergycost |  -.0350945   .1449383    -0.24   0.809    -.3192591    .2490701
   logwmprice |   .6727407   .1167159     5.76   0.000     .4439085    .9015728
   logsmprice |  -.0972975   .1002223    -0.97   0.332    -.2937924    .0991974
        _cons |  -4.125233   1.802292    -2.29   0.022    -7.658791   -.5916756
------------------------------------------------------------------------------
```

Regression C

```
      Source |       SS       df       MS              Number of obs =    2586
-------------+------------------------------           F(  5,  2580) =   11.19
       Model |  9.58964706     5  1.91792941           Prob > F      =  0.0000
    Residual |   442.09027  2580  .171352818           R-squared     =  0.0212
-------------+------------------------------           Adj R-squared =  0.0193
       Total |  451.679917  2585  .174731109           Root MSE      =  .41395

------------------------------------------------------------------------------
    logadults |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
--------------+---------------------------------------------------------------
  logavincome |  -.0072711   .0063369    -1.15   0.251    -.0196971    .0051549
   meanschool |   .0328073   .0047169     6.96   0.000      .023558    .0420567
logenergycost |  -.0193469   .0423643    -0.46   0.648    -.1024183    .0637246
   logwmprice |   .0305351   .0321064     0.95   0.342    -.0324219     .093492
   logsmprice |  -.0141252   .0276592    -0.51   0.610    -.0683617    .0401114
        _cons |    .866422   .5196463     1.67   0.096     -.152544    1.885388
------------------------------------------------------------------------------
```

Regression D

```
      Source |       SS       df       MS              Number of obs =    2586
-------------+------------------------------           F(  4,  2581) =   72.20
       Model |  47.9756292     4  11.9939073           Prob > F      =  0.0000
    Residual |  428.766411  2581  .166124142           R-squared     =  0.1006
```

```
-------------+------------------------------           Adj R-squared =  0.0992
       Total |   476.74204  2585  .184426321           Root MSE      =  .40758

------------------------------------------------------------------------------
loghouseho~e |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
logavmincome |   .0005064   .0020896     0.24   0.809    -.0035911    .0046039
  meanschool |   .0735371     .00462    15.92   0.000     .0644778    .0825965
  logwmprice |    .120554    .030535     3.95   0.000     .0606784    .1804296
  logsmprice |  -.0139521   .0270037    -0.52   0.605    -.0669032     .038999
       _cons |  -.2524043    .381133    -0.66   0.508    -.9997617     .494953
------------------------------------------------------------------------------
```

References

Becker, G. (1974). *On the Relevance of the New Economics of the Family*. The American Economic Review 64.2, 317-319.

Becker, G. (1993). *Treatise on the Family*. Cambridge: Harvard University Press.

Bumpass, L., Sweet, J. A., & Cherlin A. (1991). *The Role of Cohabitation in Declining Rates of Marriage*. Journal of Marriage and Family 53.4, 913-927.

Cherlin, A. (1990). *Recent Changes in American Fertility, Marriage, and Divorce*. Annals of the American Academy of Political and Social Science 510, 145-154.

Cherlin, A. (1992). *Marriage, Divorce, Remarriage*. Revised and enlarged edition. Cambridge: Harvard University Press.

Cutright, P. (1971). *Income and Family Events: Family Income, Family Size, and Consumption*. Journal of Marriage and Family 33.1, 161-173.

Easterlin, R., Pollack, R., & Wachter M. (1980). *Population and Economic Changes in Developing Countries*. Chicago: University of Chicago Press.

Eurostat. (2011). *Marriage Indicators*. Available from http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/search_database

Eurostat. (2011). *Divorce Indicators*. Available from http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/search_database

Greenwood, J. & Guner N. (2008). *Marriage and Divorce Since World War 2: Analyzing the Role of Technological Progress on the Formation of Households*. University of Pennsylvania Scholarly Commons, Working Paper Series.

Lanjouw, P. & Ravallion M. (1994). *Poverty and Household Size*. The World Bank Policy Research Department, Working Paper.

Living Standards Measurement Study. (2005). Albania Living Standards Measurement Survey. Available from http://econ.worldbank.org/WBSITE/EXTERNAL/EXTDEC/EXTRESEARCH/EXTLSMS/0,,contentMDK:21369054~pagePK:64168445~piPK:64168309~theSitePK:3358997,00.html

Martin, C., & Juarez F. (1995). *The Impact of Women's Education on Fertility in Latin America: Searching for Explanations.* International Family Planning Perspectives 21.2, 52-57 & 80.

McGinnis, S. L. (2003). *Cohabiting, Dating, and Perceived Costs of Marriage: A Model of Marriage Entry*. Journal of Marriage and Family 65.1, 105-116.

Musgrove, P. (1980). *Household Size and Composition, Employment, and Poverty in Urban Latin America*. Economic Development of Cultural Change 28.2, 249-266.

Nelson, J. A. (1988). *Household Economies of Scale in Consumption: Theory and Evidence*. Econometrica 56.6, 1301-1314.

Oppenheimer, V. (1967). *The Interaction of Demand and Supply and its Effect on the Female Labour Force in the United States*. Population Studies 21.3, 239-259.

Rindfuss, R., & Bumpass L. (1978). *Education and Fertility: Implications for the Roles Women Occupy*. CDE Working Paper.

Rodrik, D. (2008). *The New Development Economics: We Shall Experiment but How Shall We Learn?*. John F. Kennedy School of Government, Revised Draft.

Ruggles, S. (1997). *The Rise of Divorce and Separation in the United State, 1880-1990*. Demography 34.4, 455-466.

Ruggles, S. (1997). *Reply to Oppenheimer and Preston*. Demography 34.4, 475-479.

Sevilla-Sanz, A. (2007). *Household Division of Labor and Cross-Country Differences in Household Formation Rates*. University of Oxford, Department of Economics, Discussion Paper Series.

Stevenson, B. & Wolfers J. (2007). *Marriage and Divorce: Changes and their Driving Forces*. University of Pennsylvania Scholarly Commons, Working Paper Series.

U.S. Census Bureau, Housing and Household Economic Statistics Division, Fertility & Family Statistics Branch. (2010). *Families and Living Arrangements*. Available from http://www.census.gov/population/www/socdemo/hh-fam.html

United States Department of Labor, Bureau of Labor Statistics. (2011). *Consumer Price Index*. Available from http://www.bls.gov/cpi/data.htm

Watson, T. & McLanahan S. (2011). *Marriage Meets the Joneses: Relative income, Identity, and Marital Status*. The Journal of Human Resources 46.3, 482-517.

ACADEMIC VITA of Erich Miskewicz

Erich Miskewicz
801 Southgate Dr. Apt. C3
State College, PA, 16801
erich.miskewicz@gmail.com

Education

The Pennsylvania State University *University Park, PA*
Bachelor of Science in Economics *Graduation: May 2012*
Bachelor of Science in Finance

Honors

Graduating with honors in Economics from the Schreyer Honors College
Dean's List: 7/7 semesters

Related Experience

Delaware Investments *Philadelphia, PA*
Derivative Operations Intern *May 2011 – August 2011*

- Calculated daily exposure on swap and forward accounts and issued margin calls to brokers to meet collateral obligations in accordance with ISDA agreements
- Calculated daily variation margins in futures accounts and determined securities to be posted to meet maintenance requirements
- Performed month-end reconciliations on all account balances and updated spreadsheets to reflect daily balances of collateral

Activities

Penn State Investment Association *January2010 - December 2010*
Member/Analyst

- Took part in multiple stock pitches with the Information Technology Sector developing key drivers, bear and bull scenarios, company description, and revenue and geographic breakdown
- Actively participated in weekly educations on the investment process with in-depth looks at methods of evaluating stocks and other areas of finance

Skills

Proficient in Stata and Microsoft Excel, Powerpoint, and Word