

TEPPER HONORS THESIS

The Determinants of Self-Employment in
Mexico and Household Transitions between
2002 and 2005

Name: Yaqi Niu

Advisor: Professor Rebecca Lessem

Date: 12 May 2012

Abstract

The study of self-employment has been an intriguing factor in many developing countries. However, not many empirical studies analyze self-employment on a household level. This research focuses on Mexico in 2002 and 2005 and investigates the determinants of self-employment, transitions between self-employment and the labor market, and these transitions across asset levels all from the household perspective. The results show that high school education, asset level, previous self-employment experience and household size are important determinants of self-employment. The transition between self-employment and labor market is not homogenous across asset levels, because higher asset level increases chances of self-employment. Lastly, self-employment status decreases income, suggesting that a self-employed household will earn higher income if it were to enter the labor market.

TABLE OF CONTENTS

Introduction.....	3
Literature Review.....	6
Data	7
Estimation and Results.....	10
The Determinants of Self-Employment	10
Explaining Income changes between 2002 and 2005	13
Explaining Transitions across asset Levels	18
Conclusion and policy implications	19
Bibliography	21
Appendix.....	23
I. Job description	23
II. Change in number of self-employed members per household	23
III. Regression results with assets run as a single variable	24

INTRODUCTION

The study of self-employment has been an intriguing phenomenon in many developing countries. According to a research project by Pietrobelli, Rabellotti and Aquilina (2004), the median self-employment rate was 23.13% in the 1960s and 13.08% in the 1990, but it exhibits a wide variation across countries¹. For instance, Dominican Republic, Korea and Peru underwent a continual decrease, while Ecuador, Venezuela and Indonesia showed a U-shaped pattern during these three decades.

It is difficult for economists to pinpoint the reason to explain the presence of self-employment. Some argue that it is a matter of choice, in which people utilize their creativity to realize their business ideas. Others argue that it is a form of involuntary unemployment, because the formal labor market sector failed to supply sufficient jobs to meet the work force's needs. Maloney (2007) characterizes this problem as a "thirty-year debate over its nature: disadvantaged sector of a highly segmented labor market vs. simply an enlarged, relatively voluntary micro-firm sector".

In light of these challenges, I present the case of a single country that has experienced moderate economic growth since the 1960s: Mexico. Not only are there abundant data about Mexico available, the self-employment sector also makes up a significant portion of its employment climate. "Owners of small firms constitute 20%–25% of the Mexican work force, and employees of these firms make up approximately another 10%" (Cunningham & Maloney, 1999). In addition, my research shows that about 32.8% and 29.5% out the 3,961 households surveyed are self-employed in 2002 and 2005 respectively.

Using data from the Mexican Family and Life Survey 2002 and 2005, I attempt bridge the gap between self-employment on a macro and micro level and analyze the issue of self-employment in Mexico from a household perspective. A household perspective offers valuable insights, because an individual's decision to be self-employed should be based on his family's circumstances, as well as a collective decision among all his family members, instead of him alone. In addition, "family

¹ This result was based on a sample of 64 less developed countries and 19 developed countries from 1960 through the 1990s.

values may play an important role in shaping the organization of businesses and efficiency” (Bertrand & Schoar, 2006). For instance, I found out that the number of people that are self-employed in a household is dependent on the household size. This valuable insight will be overlooked if self-employment were to be analyzed on the individual level.

In my research, I attempt to answer the following questions: What are the determinants of self-employment in Mexico? I would like to examine if these determinants are the same as previous research since I will be observing all variables on the household level, while previous research have focused on the individual level. Secondly, I would like to seek for an explanation of a household’s transition from labor market to self-employment and vice versa between 2002 and 2005. In particular, I would like to investigate if the same reasons (such as greater income prospect) govern the case for both directions. Lastly, how does this transition differ as we break households down into various levels of asset groups? Here, I would like to find out if such transitions are homogenous across asset levels. These questions may not have been addressed explicitly for the case of Mexico, especially by analyzing self-employment on a household level.

Mexican Economy

To facilitate the understanding of self-employment in Mexico, it is important to have some brief knowledge of Mexico’s economic background. This section first provides a comparison of Mexico’s economy with other countries, followed by an overview of its self-employment sector.

On the international scale, according to the rankings of the World Bank (2011), q(PPP adjusted) in 2010. To serve as a point for comparison, Brazil is 96th, United States 18th, China 118th, and Gambia 201st. Generally, Mexico experiences moderate GDP growth every year; its GDP growth is on par with Brazil, which is another giant in South America, as well as the United States, Mexico’s neighbor and a developed country. However, it grows about three times slower than China, an emerging superpower in Asia, but faster than most underdeveloped Sub-Saharan African countries. A detailed table of the above-mentioned country’s GDP growth is displayed below:

TABLE 1: GDP GROWTH OF MEXICO AS COMPARED TO OTHER COUNTRIES

GDP Growth	2002	2005
Mexico	0.8	3.2

United States	1.8	3.1
Brazil	2.7	3.2
China	9.1	11.3
Gambia	-3.25	-0.88

On a national scale, Mexico is undergoing rising GDP per capita between 2002 and 2005. Hence, we expect the income per household from our data to have increased between 2002 and 2005. According to a separate research done by Gardner (2005), the average monthly household income in Mexico is 8422 Mexican Pesos (approximately 848 US\$) in 2002. Converting this number to an annual scale, it is more than two times higher than the average monthly household income in my data set.

TABLE 2: MEXICAN GDP PER CAPITA (CURRENT US\$) BETWEEN 2000 AND 2005

Year	GDP per capita
2000	5,817
2001	6,139
2002	6,324
2003	6,740
2004	7,224
2005	7,973

Source: World Bank

Mexico is one of the most entrepreneurial countries across the globe. Self-employment rate estimates from the OECD rank Mexico at the top among the other 28 member countries in 2000 and the Global Entrepreneurship Monitor rank Mexico 4th in its listing of 41 countries in 2003 (Fairlie & Woodruff, 2005). The self-employment rate in Mexico did not exhibit any consistent trend over the decades: it increased from 17.98% in the 1960s to 19.88% in the 1970s, but decreased to 15.92% in the 1980s, and then experienced an increase again to 17.54% in the 1990s (Pietrobelli, Rabellotti, & Aquilina, 2004).² In addition, Pagan and Sanchez (2000) remarked that the “self-employment rates are about the same for both men and women” in Mexico. With reference to Pietrobelli, Rabellotti and Aquilina’s research of the 1990s, Mexico’s self-employment rate is significantly higher than that

² These percentages only correspond to the self-employment rate in the manufacturing labor force. The agricultural sector is not included in their study.

of most European countries, but is comparable to that of its South American neighbors, with Chile at 19.61%, Peru at 22.61% and Brazil at 11.10%.

The composition of self-employed jobs is substantially different from that in the United States. According to a research by Samaniego (1998), there are six main subsectors of urban self-employment in Mexico: Restaurants, financial services, grocery stores, apparel stores, maintenance services, and automobile repair shops. These subsectors account for more than 40% of all the self-employed workers in 1996. As opposed to the common American perception of self-employment as start-up firms in Silicon Valley, self-employment firms in Mexico are less technology-related and require more basic labor skills.

LITERATURE REVIEW

As mentioned in the introduction, the nature of self-employment in developing countries is a widely controversial topic among economists. To simplify this debate for the purpose of this paper, we can generally group economists into two factions: one faction believes that self-employment is “a prosperity pull”, while the other asserts that it is “an unemployment push” (Pietrobelli, Rabellotti, & Aquilina, 2004). For example, Mariano, Edwin and Maloney (2007), as well as Yamada (1996), argue that self-employment supplies “entrepreneurial talent, which in turn enhances the sector's capacity to provide competitive earnings” (Yamada, 1996). On the other hand, Harris and Todaro (1973) equates this sector with “underemployment or disguised unemployment – the disadvantaged sector of a market segmented by rigidities in the “formal” or covered sector of the economy” (Bosch & Maloney, 2006).

Despite the disparate claims, there have been numerous attempts to reconcile these two opposing viewpoints. One of the most successful attempts is to connect self-employment with economic growth. It seems that workers turn away from self-employment and join large enterprises as “income grows and markets expand during the course of economic development” (Blau, 1987). Conversely, Ranis and Stewart (1995) and Acs et al (1994) observed that a stagnant economy, coupled by failed central planning and public sectors, can fuel people’s interest in self-employment to generate jobs and income.

There are also various studies that try to link self-employment with demographic characteristics, such as age, sex, ethnic group and education (Pagan & Sánchez, 2000) (Samaniego, 1998). In Pagan's and Sánchez's research (2000), they observed that "higher levels of education are associated with a higher employment propensity, although the variable for some secondary schooling is not significant for men." A similar result can be said about age, which is a proxy for work experience. Interestingly, age accounts for "most of the observed gender differences in employment outcomes." Lastly, they found out that "household heads are more likely to be employed than non-household heads". In general, economists reach some consensus that self-employment is more prevalent among young people (House, Ikiara, & McCormick, 1993).

With regards to the transitions between labor market and self-employment, Bosh et al (2007) have noted in the case of Brazil that these transitions are "broadly pro-cyclical and highly correlated to each other". Moreover, unemployment moves counter cyclically, hence reflecting a possibility of involuntary unemployment. Günther and Launov (2011) estimated that the "estimated size of involuntary informal employment is about 45% of the entire informal sector". As for Mexico, they observed that "the informal sector shows more constant hiring rates across the cycle, consistent with a greater degree of wage flexibility" (Bosch & Maloney, 2006).

DATA

The data that is used in this research come from Mexican Family and Life Survey (MxFLS) 2002 and 2005. The MxFLS is a "multi-thematic and longitudinal database which collects, with a single scientific tool, a wide range of information on socioeconomic indicators, demographics and health indicators on the Mexican population" (Mexican Family Life Survey, 2008). Many data, such as annual household income, annual asset value and household size, provided valuable information that is pertinent to this research.

In addition, the MxFLS is a panel dataset with data for 2002 and 2005. This is particularly useful for this research because it enables us to track the changes and transitions households undergone between these two years. To achieve this, I made use of fixed effects model to eliminate

unobserved effects, such as the IQ of each individual, and his motivation level, which may alter my results. Another point to note is with regards to outliers.

While most of the data were used directly as provided, some data were recoded to fit the purpose of this research. For the purpose of this paper, I define occupations that are self-employed as peasant on plot, family worker in a household member-owned business, boss, employer, or business proprietor, self-employed worker and finally, worker without remuneration from a business or company that is not owned by the household member. On the other hand, I define the following two types of jobs as those in the labor market: non-agricultural worker or employee, and rural laborer, or land peon (agricultural worker). Details are provided in the appendix.

TABLE 3: SUMMARY STATISTICS OF THE DATA

Variable	Mean	Std. Dev.	Min.	Max.
Household size ³	2.484	1.092	1	8
Assets02	9371.436	9629.917	7.675194	48599.33
Assets05	10732.2	10749.6	6.743675	55163.26
Income02	2356.812	2012.721	1.228031	9279.31
Income05	3292.223	2378.653	13.48735	12128.7
High School? ⁴	.181	.385	0	1

Note: 0 value has been excluded for all variables. Outliers are excluded for assets and income. Assets02 and assets05 taken at 95 percentile, income02 and income05 taken at 98 percentile. Total of 1488 observations.

In order to avoid outliers, I take assets between 0 to 95 percentile and income between 0 to 98th percentile. From the summary statistics of 1662 households above, we can discuss some general features of a Mexican household. An average Mexican household has 3 working-age members. However, all other variables (annual assets, annual income and high school education) show very high standard deviation, which indicate that the variables are scattered over a large range of values. As a result, it is difficult to describe the basic attributes of a typical Mexican household.

³ Throughout this paper, household size refers to the number of working-age members in a household.

⁴ High-School refers to whether the lead of the household has received high-school education.

To explain the data better, it will be more meaningful if we were to break down the data set into 4 groups according to households' self-employed status in both years: self-employed in 2002 and labor market in 2005, self-employed in both years, labor market in 2002 and self-employed in 2005, and finally labor market in both years. In this way, we will be able to observe if households across these four categories share certain similar attributes or disparate features. This also brings us one step closer to observe the overall transitions between labor market and self-employment between both years.

TABLE 4: SUMMARY STATISTICS OF 4 CATEGORIES

	SE02LM05	SE02SE05	LM02SE05	LM02LM05
Household size	3.681	3.747	3.614	3.120
Assets02	9371.361	11361.05	8668.92	8921.141
Assets05	10312.67	12619.69	10483.78	10285.37
Income02	2306.177	2325.887	2668.615	2576.363
Income05	3408.092	3060.304	3525.469	3610.603
High School?	.140	.151	.173	.203
Sample size	314	317	184	847
Transition rate	49.76	50.24	17.85	82.15

Note: The table above shows four possible transitions of households between 2002 and 2005: remaining in the same sector between the two years, switching from self-employed sector to the labor market sector, and vice versa. The data of rows "Household size", "Assets02", "Assets05", "Income02", "Income05" and "High School?" represent the mean for each category.

Overall, we see significant movement in and out of self-employment: approximately half the households that were in the self-employed sector remained in the same sector, while the other half moved to the labor market in 2005. About one-fifth of the households who were originally in labor market changed from the labor market to self-employment, while the remaining four-fifth remained unchanged. In other words, there is a higher tendency for self-employed households to move into the labor market than vice-versa.

We can draw three interesting observations from the table above. Firstly, households that are self-employed for both years possess the most amount of assets. This makes sense because a household needs to have a desirable amount of assets before it can start its own business. Secondly, high school education plays a more important role in households that work in labor market for both years than any other three groups. Again, this should be of no surprise since jobs in the labor market generally require higher skill sets. Lastly, we see relatively less transition from labor market to self-employment, the possible reasons of which would be discussed in the later part of this paper.

In addition to observing the general transitions between self-employment and labor market, we can further concentrate on the changes in the share of self-employed members given that the household is in the self-employed sector.

TABLE 5: PERCENTAGE OF SELF-EMPLOYED MEMBERS IN A SELF-EMPLOYED HOUSEHOLD

Year	Mean	Observations
2002	0.565	1301
2005	0.574	1170

Table 5 tells us that the percentage of self-employed members per household remains at about 56% for both years. In other words, if a household is self-employed, then more than half the working-age members of this family is self-employed.

ESTIMATION AND RESULTS

THE DETERMINANTS OF SELF-EMPLOYMENT

First, I would like to understand what factors affect self-employment. I begin with the hypothesis that asset level⁵, household size, high school education and income will affect the decision making process according to the following probit equation:

⁵ Asset level is used in replacement of annual assets because they show great variance in both 2002 and 2005

$$\begin{aligned} \text{Pr}(SE) = & \Phi(\beta_0 + \beta_1 \text{Self_Employed}_{2002?} + \beta_2 \text{Household_size} \\ & + \beta_3 \text{High_School?} + \beta_4 \text{AssetLevel0to20} + \beta_5 \text{AssetLevel20to40} \\ & + \beta_6 \text{AssetLevel40to80} + \varepsilon) \end{aligned}$$

The probit analysis is chosen because the self-employment status of a household is binary: it is either self-employed or not. In this light, it is more descriptive to use the probit model to analyze how each variable affects the probability of a household being self-employed than to use a normal regression model.

TABLE 6: PROBIT ANALYSIS OF SELF-EMPLOYMENT

Explanatory Variable	Coefficient (2002)	Coefficient (2005)	Coefficient(2005)
Asset level 0 th to 20 th percentile	-.292 *** (.068)	-.358 *** (.070)	-.295 *** (.074)
Asset level 20 th to 40 th percentile	-.233 *** (.067)	-.212 *** (.068)	-.131 * (.072)
Asset level 40 th to 60 th percentile	-.139 ** (.066)	-.063 (.067)	-.017 (.071)
Asset level 60 th to 80 th percentile	-.126 * (.066)	-.178 *** (.067)	-.1073 (.070)
Household size	.258 *** (.019)	.215 *** (.019)	.132 *** (.020)
High-School?	-.171 *** (.057)	-.140** (.058)	-.090 (.061)
Self-employed in 2002?	-	-	1.046 *** (.046)

‘***’ means significance at 0.01 level, ‘**’ at 0.05 level and ‘*’ at 0.1 level. Each regression has 3961 observations.

To better understand this table, I will use the third column as the basis of my explanation. To begin with, consider the case of five households, of which each household belongs to a different asset level. Household I has asset level 0th to 20th percentile, household II has 20th to 40th percentile and so forth. The effect of high-school will be ignored in this situation because the coefficient for high school is insignificant. In addition, the household size of each household is 3. After presenting

the value of assets each household has, I will present two different scenarios: the first scenario that the households are self-employed in the 2002 and the second scenario that they are not.

In the first scenario, there is a probability of 51.6% for household I to be self-employed, followed by 58.3% for household II. In the second scenario, there is a probability of 15.6% for household I to be self-employed and 20.05% for household II. An obvious conclusion we can draw from the above-mentioned two scenarios is that the self-employment status of a household in 2002 plays an important role in determining if the household is self-employed in 2005.

It is likewise interesting to analyze is the likelihood of a household to be in self-employment across asset levels. This point will be dealt in greater detail in the later part of this research, together with the effects of high-school education of the household lead.

Next, I would like to investigate how fraction of family members that works in family business per household (given that this household is self-employed) changes with high-school education, household size, asset level and self-employment status in 2002. I employ the following equation:

$$\begin{aligned}
 & \text{Fraction}_{self-employed\ members\ per\ household,t} \\
 & = \beta_0 \text{Household_size}_t + \beta_1 \text{High School}_t \\
 & + \beta_2 \text{Self_Employed}_{2002} + \beta_3 \text{AssetLevel0to20} + \beta_4 \text{AssetLevel20to40} \\
 & + \beta_5 \text{AssetLevel40to80} + \varepsilon
 \end{aligned}$$

TABLE 7: OLS OF FRACTION OF SELF-EMPLOYED MEMBERS PER HOUSEHOLD IN THE RESPECTIVE YEAR

Explanatory Variable	I (2002)	II (2005)	III (2005)
High-School?	-.0095 (.01492)	-.0119 (.0160)	-.0119 (.0160)
Household size	-.1604 *** (.0044)	-.1597 *** (.0048)	-.0022*** (.0193)
Asset level 0th to 20 th percentile	.0039 (.0175)	-.0022 (.0193)	-.0022 (.0193)
Asset level 20 th to 40 th percentile	-.0438 ***	-.0175	-.0175

	(.0165)	(.0177)	(.0178)
Asset level 40 th to 60 th percentile	-.0324 ** (.0158)	-.0018 (.0170)	-.0017 (.0170)
Asset level 60 th to 80 th percentile	-.0294 * (.01576)	-.0010 (.0173)	-.0010 (.0173)
Self-employed in 2002?	-	-	.0003 (.0118)

*‘***’ means significance at 0.01 level, ‘**’ at 0.05 level and ‘*’ at 0.1 level. Regression I has 1301 observations and R-squared 0.5136. Regression II has 1170 observations and R-squared 0.4838. Regression III has 1170 observations and R-squared 0.4838.*

Most variables are not significant in explaining the number of SE members in a household. In other words, the amount of assets a household has, and the annual income it earns is insufficient to explain why some members choose to work in LM while others choose not to. On the other hand, the number of working-class members per household has a significantly negative relationship with respect to the household size. In other words, every increase in household size decreases the fraction by about 10%. This may imply that the businesses have some size threshold and will not get very large. Both 2002 and 2005 showed consistent results.

EXPLAINING INCOME CHANGES BETWEEN 2002 AND 2005

In order to understand the transitions between labor market and self-employment between 2002 and 2005, it is important to analyze the determinants of a household’s annual income. However, it is possible that these determinants affect a household’s annual income to a different extent, depending on the state of employment of this household. For this reason, three different regressions are run: The first regression (first column) examines how household size, assets, and high-school education of the family lead affects self-employed households. The second regression uses the same variables, but it analyzes households in the labor market. Finally, the last regression investigates how these variables affects income after controlling for self-employment. The regression was run for the year 2002 and 2005 based on the following equation:

$$Income_t = \beta_0 Household\ size_t + \beta_1 High\ School? + \beta_2 Self_Employed_t? + \beta_3 AssetLevel0to20_t + \beta_4 AssetLevel20to40_t + \beta_5 AssetLevel40to80_t + \varepsilon$$

TABLE 8: OLS OF INCOME IN 2002

Explanatory Variable	Coefficient (self-employed)	Coefficient (labor market)	Coefficient
Household size	424.149 *** (128.330)	854.344 *** (59.308)	675.538 *** (59.890)
Asset level 0th to 20 th percentile	-2312.263 *** (502.664)	-1776.45 *** (192.482)	-1948.913 *** (207.095)
Asset level 20 th to 40 th percentile	-2291.45 *** (473.629)	-1856.66 *** (194.111)	-1996.043 *** (204.928)
Asset level 40 th to 60 th percentile	-2045.245 *** (455.957)	-1681.982 *** (195.847)	-1815.753 *** (203.754)
Asset level 60 th to 80 th percentile	-1840.634 *** (452.337)	-1260.322 *** (193.808)	-1476.543 *** (201.826)
High-School?	2681.111 *** (428.020)	1618.618 *** (154.617)	1932.934 *** (170.276)
Self-employed?	-	-	117.380 (139.444)

Note: '***' means significance at 0.01 level, '**' at 0.05 level and '*' at 0.1 level. The first column has 1301 observations and R-squared 0.0747. The second column has 2660 observations and R-squared 0.162. The last column has 3961 observations and R-squared 0.1083.

TABLE 9: OLS OF INCOME IN 2005

Explanatory Variable	Coefficient (self-employed)	Coefficient (labor market)	Coefficient
Household size	747.955 *** (91.567)	1360.727 *** (73.949)	1147.268 *** (58.253)

Asset level 0 th to 20 th percentile	-1461.142 *** (362.182)	-1743.832 *** (249.074)	-1645.125 *** (204.937)
Asset level 20 th to 40 th percentile	-1236.175 *** (332.423)	-1721.788 *** (252.493)	-1533.742 *** (202.703)
Asset level 40 th to 60 th percentile	-1319.389 *** (318.346)	-1430.467 *** (256.270)	-1369.434 *** (202.166)
Asset level 60 th to 80 th percentile	-534.4617 * (324.318)	-891.13 *** (248.735)	-761.6094 *** (199.236)
High-School?	2248.746 *** (300.867)	2521.558 *** (202.886)	2461.912 *** (168.905)
Self-employed?	-	-	-668.765 *** (139.996)

*Note: '***' means significance at 0.01 level, '**' at 0.05 level and '*' at 0.1 level. The first column has 1170 observations and R-squared 0.1332. The second column has 2791 observations and R-squared 0.1849. The last column has 3961 observations and R-squared 0.1667.*

According to the results tables, all variables are significant in explaining the amount of income earned per household. As expected, the higher the asset level, the higher the income. For instance, let us consider the income of household I to V (same as the previous section) based on the third regression in Table 8. Assuming that all five households received high-school education, then household I to V's annual income will be approximately 2315, 2430, 2590, 3200 and 3960 respectively. Also, having high-school education, which has the largest effect among all other variables in this regression, increases income by nearly 2000 pesos.

Another observation is that the household size has a greater effect on the income of households in the labor market. This may be explained by the fact that the average size for a self-employed household is 3.7, while that of a household in labor market is 3.1. As a result, any change in the size of a self-employed household will have a lower effect as compared to that in the labor market.

Lastly, self-employment status decreases income, suggesting that a self-employed household will earn higher income if it were to enter the labor market. This might provide a hint to

my third research question of why household switch sectors, because there is greater income prospect (an approximate increase of nearly 700 pesos per annum) after entering the labor market.

After investigating how assets, household size, employment status and high school education influence annual income, my next step would be to examine how changes in income between 2002 and 2005 may be attributed to changes in employment status and asset level. To investigate these changes, I made use of fixed-effects regression, which controls the ability of household members and focuses specifically on households that switched sectors. We disregard high school because the high school education of the household lead remains unchanged between 2002 and 2005. The following equation is used for the regression:

$$\begin{aligned} \Delta Income_{2002, 2005} &= \beta_0 \Delta Self_employed?_{2002, 2005} + \beta_1 \Delta Assets_{2002, 2005, self_employed} \\ &+ \beta_1 \Delta Assets_{2002, 2005, labor\ market} + \varepsilon \end{aligned}$$

TABLE 10: OLS OF CHANGE IN INCOME

Explanatory Variable	I (Household is self- employed in 2002)	II (Household is in labor market in 2002)	III (All households)
Change in self-employment status between 2002 and 2005	-277.43 *** (130.46)	13.73 (118.12)	-123.54 (87.27)
Difference in assets between 2002 and 2005 if self-employed in 2005	.015* (.008)	.004 (.0104)	.011 * (.006)
Difference in assets between 2002 and 2005 if labor market in 2005	.017 ** (.008)	.012 *** (.004)	.013*** (.004)
Self-employed in 2002?	-	-	-303.87 *** (94.60)

‘***’ means significance at 0.01 level, ‘**’ at 0.05 level and ‘*’ at 0.1 level.

For both cases, the asset level plays a major role in determining the difference in income of a household between 2002 and 2005. For regression I, the change in self-employment status can either be 0 (self-employed in both years) or -1 (self-employed in 2002 and labor market in 2005). Since the coefficient is negative, we can infer that switching from self-employment to labor market increases household's annual income. On the other hand, from regression II, the transition from labor market to self-employment cannot explain the difference in income of a household in the labor market. Lastly, we see the greatest increase in assets happens to a household shifts from self-employment to the labor market. Again, this indicates greater increase in wealth after the switch.

A subsequent question that arises is: which group experienced the greatest change in income between 2002 and 2005? To tackle this question, we need to first break down all the households into the 4 categories as mentioned in the Data Section, i.e. households which did not switch sectors for both years (SE02SE05, LM02LM05), households that transferred from the self-employment sector in 2002 to the labor in 2005 (SE02LM05), and vice-versa (LM02SE05). This gives us the following equation:

$$\Delta Income_{2002,2005} = \beta_0 SE02SE05 + \beta_1 SE02LM05 + \beta_2 LM02SE05 + \varepsilon$$

TABLE 11: OLS OF CHANGES IN INCOME WITH RESPECT TO CHANGES IN SELF-EMPLOYMENT STATUS

Remained in Self-employment in 2002 and 2005	-856.6619 *** (85.24935)
Change from Self-employment in 2002 to Labor Market in 2005	1595.463 *** (77.98568)
Change from Labor Market in 2002 to Self-employment in 2005	872.0803 *** (76.90147)
Difference in assets between 2002 and 2005	.0102036 *** (.0032832)

*Note: There are a total of 3,468 observations and R-squared 0.035. '***' means significance at 0.01 level, '**' at 0.05 level and '*' at 0.1 level.*

There is a negative correlation for households that stayed in the self-employment sector for both years, indicating a decrease in income from 2002 to 2005. On the other hand, there was a significant increase in income for households that move from self-employment to the labor market. The regression results might have provided some insights to the idea of involuntary unemployment, though they are not conclusive. We observe a very sharp increase in income should households leave the self-employment sector and enter the labor market. This drastic increase is not present in any of the other groups.

EXPLAINING TRANSITIONS ACROSS ASSET LEVELS

It should be of no surprise that the chances of households being in the self-employed sector or labor market are inhomogeneous across asset levels. In this section, I extend my probit equation under the “determinants of self-employment” section to further investigate how asset level affects a household’s probability to be self-employed.

To analyze this difference in probability in greater detail, households are divided into five groups based on their asset levels (0th to 20th percentile, 20th to 40th percentile, 40th to 60th percentile, 60th to 80th percentile, and finally 80th to 100th percentile). Using a probit function, we try to determine how different asset level affects households’ status each year.

$$\Pr(SE) = \Phi(\beta_0 + \beta_1 Household_Size + \beta_2 High_School?) + \beta_3 AssetLevel0to20 + \beta_4 AssetLevel20to40 + \beta_5 AssetLevel40to80$$

TABLE 12 PROBIT ANALYSIS OF SELF-EMPLOYMENT

Explanatory Variable	I (2002)	II (2005)
Household size	.258 *** (.019)	.215 *** (.019)
High School?	-.171 *** (.057)	-.140 ** (.058)
Asset level 0 th to 20 th percentile	-.292 *** (.068)	-.358 *** (.070)
Asset level 20 th to 40 th percentile	-.233 *** (.067)	-.212 *** (.068)

Asset level 40 th to 60 th percentile	-0.139 ** (.066)	-0.063 (.067)
Asset level 60 th to 80 th percentile	-0.126 * (.066)	-0.178 *** (.067)

*Note: Column one shows the result of the probit function in 2002, and column two for 2005. The coefficient becomes less negative as asset level increases, and this result is consistent for both years. Each regression has 3961 observations. '***' means significance at 0.01 level, '**' at 0.05 level and '*' at 0.1 level.*

As the both regressions show similar results, I will focus my analysis on regression 1. Looking at regression 1, we can conclude that the probability of self-employment differs across asset levels. Consider the five households again: each household with three working-age members and the lead of the household has high-school education. Using coefficients from regression I for calculation, the chance to be self-employed for household I to V will be 28.7%, 33.0%, 36.3%, 37.1% and 42.6% respectively.

By the same token, if we were to consider the same households, but the lead of the household has no high-school education, then the probability of self-employment will change to 37.4%, 39.3%, 42.8%, 43.2% and 48.2% respectively. We can see clearly from the regression results that the probability of self-employment increases as asset level increases. On the contrary, if the lead of the household went through high-school education, the probability of self-employment declines. This is consistent with previous researches by Pagan & Sánchez (2000) and Samaniego (1998).

CONCLUSION AND POLICY IMPLICATIONS

This paper offers numerous new insights to self-employment in Mexico from a household perspective. Firstly, probability of self-employment increases as asset level increases. Secondly, if the lead of the household went through high-school education, the probability of self-employment declines. Thirdly, the fraction of self-employed members is negatively correlated with household size, suggesting a threshold to the number of members in the self-employed business. In terms of income, high school education has a greater effect on the income of self-employed households and household size has a greater effect on the income of households in the labor market. Lastly, there

was a significant increase in income for households that move from self-employment to the labor market and may be related to the idea of involuntary unemployment.

One policy implication of this research is to focus on education policy. As seen in this research, having high-school education increases the probability to join the labor market, as well as the annual income of the household. Hence, providing more education centers and making education accessible and affordable for Mexican citizens will certainly increase their job and income prospect.

BIBLIOGRAPHY

- Acs, Z. J., Audretsch, D. B., & Evans, D. S. (1994). Why Does the Self-Employment Rate Vary Across Countries Over Time? *CPER Discussion Paper*, no. 871.
- Bertrand, M. B., & Schoar, A. (2006). The Role of Family in Family Firms. *The Journal of Economic Perspectives*, Vol. 20, No. 2, pp. 73-96.
- Blau, D. M. (1987). A Time-Series Analysis of Self-Employment in the United States. *Journal of the Political Economy*, Vol. 95, pp. 445-467.
- Bosch, M., & Maloney, W. (2006, April 1). *Gross Worker Flows in the Presence of Informal Labor Markets: Evidence from Mexico, 1987-2002*. Retrieved April 7, 2012, from Work Bank Policy Research Working Paper No. 3883: <http://ssrn.com/abstract=923230>
- Bosch, M., Goni, E., & Maloney, W. (2007, October 1). *The Determinants of Rising Informality in Brazil : Evidence from Gross Worker Flows*. Retrieved April 7, 2012, from Policy Research working paper no. WPS 4375: <http://go.worldbank.org/WDY529P3N0>
- Cunningham, W., & Maloney, W. F. (1999, September 8). *Heterogeneity among Mexico's Micro-Enterprises: An Application of Factor and Cluster Analysis*. Retrieved April 7, 2012, from World Bank Policy Research Working Paper No. 1999: <http://ssrn.com/abstract=629109>
- Fairlie, R., & Woodruff, C. (2005). Mexican Entrepreneurship: A Comparison of Self-Employment in Mexico and the United States. *NBER WORKING PAPER SERIES*, Working Paper 11527.
- Gardner, P. (2005). *Household Income and Expenditure Statistics (HIES): A comparison of eight countries*. Retrieved April 8, 2012, from <http://www.ilo.org/public/english/bureau/stat/download/articles/2006-21.pdf>
- Günther, I., & Launov, A. (2011). Informal employment in developing countries: Opportunity of last resort? *Journal of Development Economics*, 97 (1), pp. 88-98.
- Harris, J. R. (1970). Migration Unemployment and Development: A Two-Sector Analysis. *American Economic Review*, 60(1): 126-142.
- House, W. J., Ikiara, G. K., & McCormick, D. (1993). Urban self-employment in Kenya: Panacea or viable strategy? *World Development*, 21: 1205-1223.

- Maloney, W. F. (1999, April 30). *Self-employment and labor turnover*. Retrieved April 7, 2012, from Policy, Research working paper no. WPS 2102: <http://go.worldbank.org/63EPZ3QVR0>
- Maloney, W. F. (2003, January). *Informality Revisited*. Retrieved April 7, 2012, from World Bank Policy Research Working Paper 2965: <http://ssrn.com/abstract=636327>
- Mazumdar, D. (1974, October). *The Urban Informal Sector*. Retrieved April 7, 2012, from www.wds.worldbank.org
- Mexican Family Life Survey. (2008). *Mexican Family Life Survey*. Retrieved April 12, 2012, from <http://www.ennvih-mxfls.org/en/mxfls.php?seccion=1&subseccion=1&session=>
- Pagan, J. A., & Sánchez, S. M. (2000). Gender Differences in Labor Market Decisions: Evidence from Rural Mexico. *Economic Development and Cultural Change*, Vol 48, pp. 619-637.
- Pietrobelli, C., Rabellotti, R., & Aquilina, M. (2004, June). *An Empirical Study of the Determinants of Self-Employment in Developing Countries*. Retrieved April 7, 2012, from Luis Lab on European Economics: <http://www.luiss.it/ricerca/centri/llee>
- Ranis, G., & Stewart, F. (1995). V-Goods and the Role of the Urban Informal Sector in Development. *L.d'Agliano Development Studies Working Papers*, No. 85.
- Samaniego, N. (1998, April 23). *Urban Self-employment in Mexico: Recent Trends and Policies*. Retrieved May 4, 2012, from <http://www.ciln.mcmaster.ca/papers/seconf/mexico.pdf>
- The World Bank. (2011, July 1). *Gross national income per capita 2010, Atlas method and PPP*. Retrieved April 8, 2012, from <http://siteresources.worldbank.org/DATASTATISTICS/Resources/GNIPC.pdf>
- Tokman, V. E. (1978). An Exploration into the Nature of the Informal-Formal Sector Relationship. *World Development*, 6(9/10):1065-1075.
- Yamada, G. (1996). Urban Informal Employment and Self-Employment in Developing Countries: Theory and Evidence. *Economic Development and Cultural Change*, Vol 44., No 2, pp. 289-314.

APPENDIX

I. JOB DESCRIPTION

The data provided seven general job types:

1. Peasant on your plot
2. Family worker in a household member-owned business, without remuneration
3. Non-agricultural worker or employee
4. Rural laborer, or land peon (agricultural worker)
5. Boss, employer, or business proprietor
6. Self-employed worker (with or without non-remunerated worker)
7. Worker without remuneration from a business or company that is not owned by the HHm was reorganized in the following manner.

Job type was reorganized such that job type 2 and 7 belong to family business and are not paid, 3 and 4 belong to labor market, and finally, 1, 5 and 6 belong to self-employment.

II. CHANGE IN NUMBER OF SELF-EMPLOYED MEMBERS PER HOUSEHOLD

Change	Frequency	Percentage	Cumulative
-3	8	0.62	0.62
-2	26	2.00	2.62
1	168	12.92	15.54
0	916	70.46	86.00
1	155	11.92	97.92
2	21	1.62	99.54
3	4	0.31	99.85
4	1	0.08	99.92
6	1	0.08	100.00
Total	1300 (Corr: 0.325)		

A few points could be made here. Firstly, about 40% of a household, which is involved in self-employment, is self-employment. This means that, if there is a family of 5, then there will be 2

members who are self-employed. Moreover, this probability stayed almost the same between 2002 and 2005. Secondly, we notice that about 70% of the households have the same number of self-employed family members in 2002 and 2005.

III. REGRESSION RESULTS WITH ASSETS RUN AS A SINGLE VARIABLE

$$\text{Pr}(SE) = \Phi(\beta_0 + \beta_1 \text{Self_Employed}_{2002} + \beta_2 \text{Household_size} + \beta_3 \text{Assets} + \varepsilon)$$

PROBIT ANALYSIS OF SELF-EMPLOYMENT (IN COMPARISON TO TABLE 6)

Explanatory Variable	Coefficient (2002)	Coefficient (2005)	Coefficient(2005)
Assets	6.82e-06 *** (2.14e-06)	2.47e-06 (1.98e-06)	1.09e-06 (2.08e-06)
Household size	.257 *** (.019)	.211*** (.019)	.130 *** (.021)
High-School?	-.159*** (.061)	-.139 ** (.062)	-.089 (.066)
Self-employed in 2002?	-	-	1.021*** (.048)

*** means significance at 0.01 level, ** at 0.05 level and * at 0.1 level.

To better understand the table, let us analyze the situation of a median family in this sample with family size 3, asset of 9443.651 in 2002 and 10,732 in 2005. If this family is self-employed in 2002 and the household head has high school education, then there is 54% chance that this family will remained self-employed in 2005. If the family head does not have a high school education, there is a higher probability of almost 58% that this family will remain self-employed in 2005. On the other hand, if the family head has high school education but worked in the labor market in 2002, there is a mere 18% chance that his family will enter the self-employment sector in 2005. Likewise, a family that has worked in the labor market in 2002 and without high school education has about 21% chance to be self-employed in 2005. This shows that the likelihood for an average-income and household size family to be self-employed in 2005 depends greatly on whether this family is self-employed in 2002.

$Fraction_{self-employed\ members\ per\ household,t}$

$$= \beta_0 Household\ size_t + \beta_1 Assets_t + \beta_2 High\ School_t + \beta_3 SE + \varepsilon$$

OLS OF FRACTION OF SELF-EMPLOYED MEMBERS PER HOUSEHOLD IN THE RESPECTIVE YEAR (IN COMPARISON TO TABLE 7)

Explanatory Variable	Coefficient (2002)	Coefficient (2005)	Coefficient (2005)
High-School?	-.011 (.011)	-.029** (.013)	-.029 ** (.013)
Assets	1.59e-07 (3.52e-07)	-3.19e-07 (3.98e-07)	3.26e-07 (3.98e-07)
Household size	-.101 *** (.002)	-.097 *** (.002)	-.097 *** (.002)
Self-employed in 2002?	-	-	.005 (.008)

'***' means significance at 0.01 level, '**' at 0.05 level and '*' at 0.1 level.

$$Income_t = \beta_0 Household\ size_t + \beta_1 \Delta Assets_t + \beta_2 High_School? + \beta_0 Self_Employed_t? + \varepsilon$$

OLS IN 2002 (IN COMPARISON TO TABLE 8)

Explanatory Variable	Coefficient (self-employed)	Coefficient (labor market)	Coefficient
Household size	212.960 *** (26.139)	352.660 *** (16.293)	308.217 *** (13.839)
Assets	.030 *** (.004)	.034 *** (.003)	.033 *** (.002)
High-School?	1349.14 *** (120.504)	986.771 *** (66.208)	1088.384 *** (58.380)
Self-employed?	-	-	207.829 *** (47.656)

Note: '***' means significance at 0.01 level, '**' at 0.05 level and '*' at 0.1 level. The first column has 1179 observations and R-squared 0.0945. The second column has 2466

observations and R-squared 0.15. The last column has 3645 observations and R-squared 0.1291.

OLS IN 2005 (IN COMPARISON TO TABLE 9)

Explanatory Variable	Coefficient (self-employed)	Coefficient (labor market)	Coefficient
Household size	438.883 *** (33.34444)	626.850 *** (18.172)	575.712 *** (16.011)
Assets	.033 *** (.005)	.045 *** (.003)	.042 *** (.002)
High-School?	1211.218 *** (156.350)	494.122 *** (70.746)	638.201 *** (65.153)
Self-employed?	-	-	-222.108 *** (58.019)

Note: '***' means significance at 0.01 level, '**' at 0.05 level and '*' at 0.1 level. The first column has 1044 observations and R-squared 0.1355. The second column has 2593 observations and R-squared 0.2301. The last column has 3637 observations and R-squared 0.1959.

All variables are significant in explaining the amount of income earned per household. Interestingly, high school education has a greater effect on the income of self-employed households. This is counterintuitive, because previous results showed that having high school education decreases the probability of households remaining in the self-employed sector. It may be interpreted as the desire of a self-employed household to switch to the labor market is greater than the prospect of higher income for staying in self-employment.

Another observation is that the household size has a greater effect on the income of households in the labor market. This may be explained by the fact that the average size for a self-employed household is 3.7, while that of a household in labor market is 3.1. As a result, any change in the size of a self-employed household will have a lower effect as compared to that in the labor market. Assuming that this is the case, let us consider two average households in the year 2002. The first household is self-employed, with assets worth 10,207, household size of 2.78, and the family

lead has high school education. According to the first regression, this household will have an income of approximately 2250 pesos. The second household works in the labor market, has asset worth 9,596, household size of 2.32, and the family lead also has high school education. According to the first regression, this household will have an income of approximately 2140 pesos. We see that an average household in either sector will have approximately the same annual income. By the same token, if the first household were to work in the labor market, this household's annual income will increase to 2323 pesos, which is not a significant increase.