Two Applied Economics Essays: Trade Duration in U.S. Fresh Fruit and Vegetable Imports & Goods-Time Elasticity of Substitution in Household Food Production for SNAP participants and nonparticipants

Jeta Rudi

Thesis submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Master of Science
In
Agricultural and Applied Economics

Jason H. Grant, Co-chairperson
Wen You, Co-chairperson
George C. Davis
Everett B. Peterson

July 6, 2012
Blacksburg, VA

Keywords: Trade duration, Fresh fruits and vegetables, Survival analysis, Elasticity of substitution, Household food production, Supplemental Nutrition Assistance Program

Two Applied Economics Essays: Trade Duration in U.S. Fresh Fruit and Vegetable Imports & Goods-Time Elasticity of Substitution in Household Food Production for SNAP participants and nonparticipants

Jeta Rudi

ABSTRACT

The first study investigates the factors that impact the duration of U.S. fresh fruit and vegetable imports. We employ both survival analysis (Kaplan Meier estimates and Cox proportional hazards model) as well as count data models. Our results indicate that SPS treatment requirements positively impact the duration of trade while new market access has the opposite effect. Other factors typically included in trade duration models (such as: GDP, transportation costs, tariff rates, etc.) were also investigated. We also employ a probit model to understand the factors impacting the probability that a country selects into exporting fresh fruits and vegetables to the United States.

The second study estimates the goods-time elasticity of substitution for Food Stamp/SNAP participants versus non participants. We find that the elasticity of substitution for SNAP participants is not statistically different from zero. This indicates that SNAP participants have Leontief production function in household food production, implying that increasing the amount of SNAP benefits paid to participants will not lead to more food production if the time households dedicate to food preparation remains unchanged. This finding extends the analysis done by Baral, Davis and You (2011) and offers insights for policies related to the SNAP program.

DEDICATION

To my family – with much love

&

To the memory of John G. Fee – his profound vision changed the course of my life

ACKNOWLEDGEMENTS

During the past two years at Virginia Tech, I have learned so much, and yet I feel as if I have more unanswered questions now than when I first arrived. This journey of learning and quest to understand Economics was made possible primarily by the members of my Committee. They have been my professors and mentors, and I am deeply grateful for the knowledge, passion and the sense of curiosity that they have transmitted to me. Dr. Peterson and Dr. Grant have challenged me to explore a field I had little knowledge of, while providing invaluable direction and expertise. I appreciate their patience in answering even my most basic questions, as well as their guidance towards the literature in order to explore and learn things on my own. Dr. You has taught me the importance of rigorous econometric analysis and attention to detail. Her knowledge of the field and the dedication to her work have provided both inspiration to work hard as well as aspiration to achieve a lot. Being Dr. Davis' student has been an integral part of my education at Virginia Tech. He has taught me that research ideas are born in a curious mind, and that it is not skills that should dictate what research question to seek to answer, but rather that research ideas should guide which skills and set of tools need to be used. He has taken the time to provide me with important advice on my education, career, and the path to becoming a respectable economist and researcher – for all of these, I will always be deeply grateful.

Other faculty members in the Department have also contributed enormously to my learning and academic success, especially, Dr. Taylor, Dr. Mills, Dr. Norton and Dr. Stephenson. I am also grateful to Dr. Nicolai V. Kuminoff for the privilege of working with him, and for his invaluable advice on many aspects of my education. During the two years at Virginia Tech, I have not only had the chance to broaden my knowledge in the field of Economics, but also in a lot of other aspects. My friends and colleagues have given a great contribution in that regard, while also giving me the strength and inspiration to continue on even in the most difficult times. I would especially like to thank Nadezda Amaya Urquieta, Keeletlhoko Faith Badirwang, Vanessa Viñoles Gómez, Vanessa Carrion, Mina Hejazi, Maria Guadalupe Bravo-Vinaja, Gertrude Nakakeeto, Olivera Jankovska and Annah Latane.

My first Economics professor, Dr. Jonathan Lipow, has instilled in me the passion for Economics that I only hope to someday be able to instill in my students. I am very grateful to him for introducing me to this curious and important field. Throughout the four years at Berea College, Dr. Caryn Vazzana has been my professor, my mentor and my adviser. She has given me a strong foundation of the field in her classes, while also providing guidance in various research projects. The success of all her students speaks to her amazing achievements as a professor and mentor. I am lucky to be her student. I am also grateful to Berea College, for providing me with an opportunity to seek my dream of obtaining a good quality education. My many friends at Berea have been the most integral part of my life and education in the United States; I am thankful to each and every one of them for the privilege of being their friend.

I am also very grateful to the families of Mr. Besnik Gojani and Mr. Sherbet Vula in Kentucky and the family of Dr. Pavli Mykerezi in Virginia for providing me with a home away from home. And finally, I am eternally grateful to my family. My parents' vision, encouragement and selfless love are the most important factors that have led to my success until now and any future achievements. My brother, Nderim Rudi, has inspired me, challenged me, provided me with comfort and advised me – each in crucial moments. My sisters, Dita, Erza and Brittani Rudi have always been very loving and supportive; I am blessed to have them in my life. And I owe the most beautiful and joyful moments in the past two years to my nephews, Mir Pula and Van Ilirian Rudi. They are a source of strength and joy for our whole family.

TABLE OF CONTENTS

DEDICATION ... iii
ACKNOWLEDGEMENTS ... iv
LIST OF FIGURES ... viii
LIST OF TABLES ... ix
Chapter 1 - Survival of the Fittest: Explaining Export Duration and Export Failure in the U.S. Fresh Fruit and Vegetable Market ... 1
1.1. Introduction ... 1
1.1.1. Objectives ... 7
1.1.2. Organization ... 8
1.2. Literature Review on Trade Duration ... 9
1.3. Empirical Methods ... 13
1.3.1. Survival Analysis ... 13
1.3.2. Count Data Model ... 19
1.3.3. Probit Model ... 21
1.3.4. Data ... 23
1.4. Results ... 24
1.4.1. Survival Analysis ... 24
1.4.2. Count Data Model ... 32
1.4.3. Probit Model ... 37
1.5. Conclusions ... 40
References ... 44
Appendix 1.A. ... 47
Appendix 1.B. ... 48
Chapter 2 - Goods-Time Elasticity of Substitution in Household Food Production for SNAP participants and nonparticipants ... 50
2.1. Introduction ... 50
2.1.1. Objectives ... 51
2.1.2. Organization ... 52
2.2. Literature Review ... 52
2.2.1. Goods-Time Elasticity of Substitution ... 52
2.2.2. Supplemental Nutrition Assistance Program Participation ... 53
2.3. Conceptual and Empirical Framework ... 56
2.4. Data ... 60
2.5. Empirical Results ... 67

2.6. Conclusions 78
References 80

LIST OF FIGURES

Figure 1. 1. Survival Function by *Treat* 25
Figure 1. 2. Survival Function by *NMA* 25

Figure 2. 1. Radius Matching (Caliper 0.01) 76
Figure 2. 2. Radius Matching (Caliper 0.001) 76

LIST OF TABLES

Table 1. 1. U.S. Imports of Fresh Fruits, 2006-2008 4
Table 1. 2. U.S. Imports of Fresh Vegetables, 1996-2008 5
Table 1. 3. Trade Duration by Country for APPLES, 1996-2008 6
Table 1. 4. Cox Proportional Hazards Model Results 27
Table 1. 5. Count Model (Poisson) Results 33
Table 1. 6. Probit Model Results 38

Table A. 1. List of Exporting Countries of FF&V 47

Table B. 1. Explanatory Variables' Description 48

Table 2. 1. Summary Statistics 65
Table 2. 2. Goods-Time Elasticity of Substitution Results Before Propensity Score Matching .. 69
Table 2. 3. Summary Statistics After Propensity Score Matching 73
Table 2. 4. Goods-Time Elasticity of Substitution Results After Propensity Score Matching..... 77

Chapter 1 - Survival of the Fittest: Explaining Export Duration and Export Failure in the U.S. Fresh Fruit and Vegetable Market

1.1. Introduction

In the last twenty years, U.S. imports of fresh fruits and vegetables (FF&V) have increased sharply from $2.3 billion in 1989 to $14.4 billion in 2011.[1] Factors that have contributed to this increase include increasing consumer incomes, changes in consumer preferences for year round supply of fresh produce, a greater variety of fresh fruits and vegetables, reductions in tariffs and non-tariff barriers as well as increased market access through bi- and multi-lateral trade agreements (Aksoy and Beghin, 2005; Clemens, 2004; Lucier *et al.*, 2006; Johnson, 2010). Recent work by Besedeš and Prusa (2006a, 2006b) has shown that the duration of trade relationships tends to be short with numerous entries and exists (leading to multiple spells of service) in a market. This suggests that the growth in trade may be occurring more through increases in the intensive margin (e.g., increasing value of trade for existing trade relationships) rather than the extensive margin (e.g., increases in trade in new products or to new countries).

The duration of fresh fruits and vegetable exports to the U.S. also exhibit a similar pattern. Defining relationship duration as the number of years that a country exports a specific fresh fruit or vegetable commodity during our sample period, 1996-2008, or a 13 year timeframe, the average duration was 6.1 years. Duration for fresh fruits was slightly longer at 6.3 years compared to fresh vegetables at 5.9 years. In addition, multiple spells of service, defined as a

[1] The statistics are based on author's calculations from data retrieved from the U.S. Department of Agriculture – Foreign Agricultural Service (USDA – FAS) and the United Nations Commodity Trade Statistics Database (UN Comtrade).

country that stops exporting a given fresh fruit or vegetable and then begins to export again at a later date, are also common.[2] Between 1996 and 2008, the average number of spells per country-commodity pair was 1.45, with a maximum of 5 spells. The length of spells varies across country-commodity pairs. On average, spells of service last for 4.2 years each. Spells of service for fruits have a longer average length (4.5 years) compared to those for fresh vegetables (4.0 years).

As shown in Tables 1.1 and 1.2, there is significant variation in the average spell length across fresh fruits and vegetable products.[3] For fresh fruits, average spell length ranges from 2.6 years for apricots to 6.4 years for bananas. For fresh vegetables, average spell length ranges from 2.0 years for potatoes to 5.9 years for broccoli. In addition, even for a commodity that has a relatively long average spell length, there is variability in duration and the number of spells between exporting countries. For example, consider the case of apples that has an average spell length of 5.3 years. Of the 16 countries that exported apples to the U.S. between 1996 and 2008, only 5 exported at least $10,000 per year throughout the entire period (See Table 1.3). Hence, those five countries (Argentina, Canada, Chile, Japan and New Zealand) had one spell of service each, which lasted for 13 years. Eight exporting countries had one spell of service each that lasted between 1 and 5 years. Finally, 3 exporting countries had multiple spells of service. Note that as reported in Tables 1.1 and 1.2, across all commodities, the average value of exports is positively correlated with the average length of spells (0.48) and negatively correlated with the

[2] Besedeš and Prusa (2006a) provide the following definition for a spell of service and spell length in the trade literature: "If the United States imports product i from country c from 1976 to 1980, the ci^{th} trade relationship has a spell length of five" since exports have continued for five consecutive years. Multiple spells of service occur when each spell of service ends with an exit from the market, only to enter back after one or more years. Note the difference between spell length defined above, and relationship duration (which is the number of years across all spells of service, in cases of multiple spells).

[3] The data used here are similar to Karov *et al.* (2009) and Jankovska *et al.* (2011) who considered FF&V exports from 89 countries. Appendix A shows the list of the 89 exporting countries.

average number of spells (-0.27). This implies that longer spells of service are correlated with higher trade volume while multiple spells of service correlated with smaller trade volume.

Table 1. 1. U.S. Imports of Fresh Fruits, 2006-2008

	Fresh Fruits	No. of Countries Exporting Commodity	Average Export Value ($ mil.)	Average Number of Spells per Relationship	Average Length of Spells
1	Apples	16	14.370	1.3	5.3
2	Apricots	12	1.018	1.6	2.6
3	Avocados	11	43.057	1.2	5.2
4	Bananas	29	67.161	1.2	6.4
5	Cherries	16	4.011	1.4	2.8
6	Cranberries & Blueberries	18	13.680	1.4	3.9
7	Currants	7	0.129	1.9	2.2
8	Grapefruit	9	0.559	1.2	3.7
9	Grapes	15	75.478	1.3	6.1
10	Kiwifruit	12	6.544	1.6	4.0
11	Lemons	28	1.423	1.9	3.0
12	Limes	17	9.139	1.6	4.0
13	Mandarins & Clementines	18	9.894	1.4	5.4
14	Mangoes	23	11.569	1.4	5.5
15	Melon	23	13.771	1.4	6.0
16	Oranges	22	5.191	1.4	4.6
17	Papayas	18	7.031	1.4	4.5
18	Peaches & Nectarines	13	9.736	1.3	4.0
19	Pears & Quinces	17	8.764	1.1	6.1
20	Pineapples	28	14.304	1.5	4.7
21	Plums & Sloes	23	4.491	1.4	3.1
22	Raspberries & Blackberries	13	8.900	1.5	4.3
23	Strawberries	18	9.907	1.5	3.3
24	Watermelons	14	10.827	1.2	6.0

(UN Comtrade, 2010)

Note: The statistics only include trade relationships for which the customs value is greater than or equal to 10,000 USD.

Table 1. 2. U.S. Imports of Fresh Vegetables, 1996-2008

	Fresh Vegetables	No. of Countries Exporting Commodity	Average Export Value ($ mil.)	Average Number of Spells per Relationship	Average Length of Spells
1	Asparagus	20	14.658	1.4	5.4
2	Broccoli	6	8.251	1.3	5.9
3	Brussels Sprouts	7	2.126	1.1	5.1
4	Cabbage	16	2.681	1.3	3.2
5	Carrots	8	6.344	1.8	4.1
6	Cauliflower	11	1.464	1.5	3.5
7	Cucumbers	19	26.889	1.5	4.4
8	Eggplants	20	4.600	1.3	3.8
9	Fresh Beans	26	3.535	1.4	3.4
10	Garlic	35	4.363	1.5	3.4
11	Globe Artichoke	11	0.504	1.3	3.0
12	Head Lettuce	10	5.130	1.7	3.6
13	Jicamas, Pumpkins, Breadfruit	19	1.954	1.3	5.0
14	Leaf Lettuce	10	3.129	1.6	3.9
15	Leeks	11	0.985	1.3	4.8
16	Mushrooms And Truffles	38	2.909	1.6	3.9
17	Okra	14	2.596	1.8	2.6
18	Onions	34	9.333	1.8	4.0
19	Peppers	36	30.639	1.4	4.6
20	Potatoes	15	31.833	1.5	2.0
21	Spinach	7	2.027	1.3	3.7
22	Squash	21	14.183	1.6	4.2
23	Tomatoes	20	89.353	1.5	5.0

(UN Comtrade, 2010)

Note: The statistics only include trade relationships for which the customs value is greater than or equal to 10,000 USD.

Table 1. 3. Trade Duration by Country for APPLES, 1996-2008

	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
Argentina	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Australia					✓	✓	✓						
Brazil	✓	✓		✓	✓			✓	✓			✓	✓
Canada	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Chile	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
China							✓	✓	✓	✓	✓		
Guatemala					✓							✓	
Japan	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Korea								✓					
Mexico									✓	✓			
New Zealand	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
South Africa	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓
Switzerland					✓								
Thailand											✓	✓	
Uruguay	✓	✓	✓	✓	✓								
Vietnam													✓

(UN Comtrade, 2010)

1.1.1. Objectives

Given the high-level of dynamism observed in U.S. fresh fruit and vegetable imports, the main objective of this paper is to identify the factors that affect the duration of trade in this market. Such factors may include both the variables that are usually included in trade duration studies as will be explained in more depth below, as well as variables related to policies that regulate the imports of FF&V. To our knowledge, this study is the first of its kind to link trade duration in fresh fruits and vegetables to phytosanitary regulations and other Animal and Plant Health Inspection Service (APHIS) related regulations for product entry.

To prevent the introduction of plant pests and diseases from the importation of fresh fruits and vegetables, APHIS has the authority to promulgate import regulations by country and commodity under the Plant Protection Act of 2000. Each exporting country must petition APHIS before being permitted to export a specific product to the United States. If the exporting country has identified pest risks and has not developed approved mitigation practices, APHIS can deny the petition. However, if mitigation measures that reduce the risk of pest or disease outbreaks can be identified, APHIS will recommend that the product be allowed to enter subject to a set of phytosanitary (SPS) measures. If no pest risks are identified, then APHIS will recommend that the product be allowed to enter subject to routine inspection requirements. In this paper, we will focus our attention on the effects of phytosanitary treatments, such a methyl bromide fumigation or cold treatment, on duration because these treatments are costly for exporters to implement.

A second policy related issue we will investigate is whether countries that gained new market access (NMA) for a specific fresh fruit or vegetable have shorter trade duration than established exporters. The Agreement on Agriculture (AOA) in the Uruguay Round of multilateral trade negotiations led to an increase in new market access, the procedure by which countries are allowed to export new products to the U.S., at times subject to certain conditions such as the SPS requirements outlined above (Jankovska *et al.*, 2011). Between 1996 and 2008, APHIS granted new import permits to 67 country-product pairs (Jankovska *et al.,* 2011).

Several different methodologies have been utilized to analyze trade duration. The most commonly used method is survival analysis (Besedeš and Prusa 2006a, 2006b). However, Hess and Persson (2010) argue that this method may not be appropriate in studies of bilateral trade for several reasons. First, many spells of service have the same length since annual trade data masks the dynamics of trade shipments which actually occur on a much more frequent basis (monthly, or even daily). Continuous time methods face difficulties in dealing with this issue and can lead to biased coefficient estimates as well as biased standard errors. It is also difficult to properly control for unobserved heterogeneity in survivor analysis and the assumption of proportional hazards may not be appropriate in all applications. As a robustness check, we will use a Cox Proportional Hazards (PH) model (survival analysis method) and a Poisson (count data) model to identify the factors that affect trade duration. However, because neither of these models can provide any insight on what factors affect the decision to export, we also use a Probit model to understand how the probability of exporting is affected by factors such as SPS treatment requirements, the costs of trading, and exporting country characteristics, etc.

1.1.2. Organization

The study is organized as follows. Section 1.2 summarizes the literature on trade duration. Section 1.3 offers a brief overview of the three methodologies considered: Survival Analysis (with emphasis on the Cox Proportional Hazards Model), Count Data (Poisson) Model, and Probit model. This section also provides details on the different methodologies used in previous studies. The empirical results from all three models are offered in section 1.4. Section 1.5 summarizes the findings and offers the conclusions, limitations and recommendations of this study.

1.2. Literature Review on Trade Duration

In their seminal work Besedeš and Prusa (2006a) found that the duration of trade of a given product between two countries is, on average, relatively short lived and there are a large number of entries and exits. They compared U.S. imports during two different periods, from 1972 to 1988 and from 1989 to 2001.[4] From 1972 to 1988 the commodities were classified based on the 7-digit Tariff Schedule (TS) system, while from 1989 to 2001 the classification used was the 10-digit Harmonized System (HS). During both periods of analysis, more than half of trade relationships fail after one year of service and about 70% fail within two years of service. The mean spell length is about 3 years but the median spell length is only 1 year, hence the distribution is positively skewed. Despite the fact that the majority of trade relationships seem to be very short lived, Besedeš and Prusa find that the conditional probability of failure decreases with duration, where failure is defined as exit from the market. More specifically, using a nonparametric approach and estimating the Kaplan Meier (KM) survivor function, the authors conclude that relationships with exporters from North America have the highest rate of survival with 78 percent of the relationships surviving the first year.[5] On the other hand, the relationships with exporters from Africa have the lowest rate of survival with only 52 percent of the relationships surviving after the first year and only 29 percent after four years of service. The results are robust across different model specifications (Besedeš and Prusa, 2006a).

[4] Data used in this as well as the rest of the studies that will be discussed in this review are all country-level data unless otherwise indicated, mostly because of the lack of widely available import/export data at the firm level.
[5] The Kaplan-Meier estimator is "a non parametric estimate of the survivor function $S(t)$, which is the probability of survival past time t or, equivalently, the probability of failing after t." (Cleves et al. 2010, p.93).

In subsequent work, Besedeš and Prusa (2006b) analyzed whether there are differences in the duration of trade between differentiated and homogeneous goods using the same system of product classification as the in their previous study. They hypothesized that import spells of differentiated goods have a higher chance of survival than import spells of homogeneous goods. They also hypothesized that trade duration is positively correlated with the size of initial transaction for both types of goods. Using the nonparametric approach (KM estimates) they find that the probability of survival to year two for spells of differentiated goods is higher (69 percent of those relationships survive) than for homogeneous goods (55 percent of those relationships survive to year two). This finding is confirmed by estimating the hazard rate through a Cox proportional hazard (PH) model.[6] They find that spells of homogeneous goods have a 23 percent higher hazard rate compared to differentiated goods. Similarly, the authors find evidence that trade relationships with a smaller initial transaction size have a lower survival rate than those with larger initial transactions, for both product types. When the sample includes only spells of service for which the initial transaction size is at least $100,000, 69 percent and 55 percent of trade relationships involving differentiated and homogeneous goods respectively, survive to year two. However, when the sample is restricted to spells of service with initial transaction size exceeding $1 million, 99 percent of relationships involving differentiated goods and 75 percent of relationships involving homogeneous goods survive to year two. The Cox PH model give similar results (Besedeš and Prusa, 2006b).

In the U.S. data for differentiated goods, only three percent of the trade relationships between the U.S. and another country that have more than $1 million in initial sales fail after one year, but almost 50 percent of trade relationships with initial sales of less than $10,000 fail by the

[6] The hazard function or hazard rate is "the (limiting) probability that the failure event (in our case, exit from the market) occurs in a given interval, conditional upon the subject (in our case, spell of service) having survived to the beginning of that interval, divided by the width of the interval." (Cleves *et al.* 2010, p.7).

second year (Besedeš, 2008).[7] One explanation for this is that relationships with higher search costs, which may include uncertainty and lack of information, will begin with lower initial sales and will be more likely to fail. Besedeš (2008) uses a set of four variables to approximate the search costs incurred for a trade relationship: distance between the U.S. and exporting country, common language, contiguity, and number of potential suppliers in the market.[8] In addition, variables measuring supplier reliability, relative cost of trading, and initial transaction size are also included in a stratified Cox PH model.[9] [10] The author finds that the lower the search costs, the lower is the hazard rate (or conversely, the higher the conditional probability of survival). For example, trade relationships with Mexico have a 28 percent lower hazard rate due to its contiguity with the U.S., while trade relationships with English speaking countries face a 3 percent lower hazard rate compared to those from non-English speaking countries. Distance on the other hand does not have a significant impact on the hazard rate, increasing the distance of the exporting country from the U.S. by 1,000 kilometers leads to a 0.015 percent increase in the hazard rate. Finally, increasing the number of potential suppliers to the U.S. decreases the hazard rate. This is a counterintuitive result because it suggests that more competition increases the conditional probability of survival in the market. The author also finds that more reliable exporters, as measured by a higher GDP per capita and a lower number of spells of service per country-commodity pair, face a lower hazard rate. The cost of trading as measured by the transportation costs and changes in the exchange rate also have the expected impact, although the impact of higher transportation costs is much lower than that of the exchange rate (Besedeš,

[7] The data classification in this study is the same as in Besedeš and Prusa (2006 a, 2006b). That is, from 1972 to 1988 the 7-digit TS system is used, whereas from 1989 to 2001 the 10-digit HS is used.

[8] Since firm characteristics are not observed, the author uses a combination of country characteristics as well as the number of other potential suppliers as proxies of the search costs incurred (Besedeš, 2008).

[9] The stratified Cox PH model allows the baseline hazards to differ by group, but the coefficients of the covariates are restricted to be the same (Cleves et al. 2010, p.144).

[10] The model was stratified by region and one-digit SITC industries (Besedeš, 2008).

2008). In subsequent studies, Besedeš also explores the impact of NAFTA on duration of trade partnerships between the U.S., Canada and Mexico (2011b), as well as offering a view on how the findings in trade duration can be incorporated on the debate of the role of extensive versus intensive margin of trade (Besedeš and Prusa, 2010).

Pursuing the same topic in another region, Nitsch (2009) conducted a similar analysis for German imports over the period of time 1995-2005. He also found that on average, trade relationships between countries for a given product tend to be short-lived with numerous entries and exits. In studying the factors that impact duration, he concluded that Gravity model variables that were successful in explaining the duration of U.S. import relationships also do well in explaining the duration of German import relationships.[11] For example, trade relationships with countries with higher GDP have lower hazard rates hence their spells of service have longer duration. Distance has a negative impact on the duration of spells. Contrary to what one would expect, spells of service from richer countries (as measured by GDP per capita) have higher hazard rates. As in the case with U.S. import relationships, spells of service from countries bordered with Germany or those that share a common language last longer. The author further finds that the impact of exchange rate is not statistically distinguishable from zero; neither does membership of the exporting country in the EU have any impact on the hazard rates (Nitsch, 2009).

In addition to Nitsch (2009), other studies on trade duration have focused on trade relationships for countries in East and West Africa (Cadot *et al*., 2011) and East Asia (Obashi, 2010). In contrast to most studies on the topic, Cadot *et al.* (2011) use firm level rather than country level data. While the results vary somewhat, relatively short durations and high churning

[11] The author employs both the KM estimates method as well as the stratified Cox PH model. He stratified the Cox model by World Bank regions and 1-digit industries (Nitsch, 2009).

among exporters remain. Additionally, Besedeš (2011a) analyzes the duration of trade partnerships among the transition countries of Central and Eastern Europe. Several studies have also explored the factors that impact trade duration for developing country exporters (Brenton *et al.* 2009, Brenton *et al.* 2010). While several authors have explored the problem of trade duration for different regions involving diverse sets of trading partners, to the best of our knowledge this study is the first to focus exclusively on a very narrow group of commodities matched to very specific non-tariff measures and other trade cost factors.

1.3. Empirical Methods

1.3.1. Survival Analysis

Survival Analysis is the method of analyzing the time until the occurrence of an event. It is widely used in the field of Medicine to understand the risk from diseases, reactions to medicine, and whether different medical treatments lead to different rates of survival among patients with a given medical condition. In Economics, survival analysis has been used to understand the factors that lead to exit from unemployment, and more recently to understand the duration in trade relationships.[12] Our analysis falls into the latter category in that we seek to understand the factors impacting the duration of U.S. FF&V imports.

Let T be a nonnegative random variable denoting the time to a failure event with a probability density function $f(t)$ and a cumulative distribution function of $F(t)$.[13] The survival function $S(t)$ is defined as:

$$S(t) = 1 - F(t) = Pr(T > t). \qquad (1.1)$$

[12] See for example Mills (2000) on how survival analysis methodology is used in unemployment literature.
[13] Failure can be defined in many different ways. In Medicine, failure can be the event of a patient dying. In the unemployment duration literature, failure is the event of exiting unemployment (e.g. finding a job). In the trade duration literature, failure occurs when an exporter stops exporting (i.e. the spell of service ends).

Thus, the survivor function is simply the reverse cumulative distribution function of T (Cleves *et al.* 2010, p.7). The hazard function, $h(t)$, is defined as the probability that the failure event occurs in a given interval, conditional upon the subject having survived to the beginning of that interval, divided by the width of the interval. Specifically:

$$h(t) = \lim_{\Delta t \to 0} \frac{Pr(t + \Delta t > T > t \mid T > t)}{\Delta t} = \frac{f(t)}{S(t)}. \tag{1.2}$$

For example, if we would like to know the hazard function for time $t = 10$ it will be: $h(t = 10) = \Pr(11 > T > 10|\mathrm{T} > 10)$. Depending on its form and the impact of specific covariates on the survival experience, the survival function can be estimated under three distributional assumptions: nonparametric, semiparametric and parametric approaches.[14]

In the nonparametric approach, no assumptions are made about the functional form of the survivor function and nonparametric tests. For example, the log-rank or Wilcoxon tests are used to determine if there are differences between subgroups of individuals in the sample. Most previous studies on trade duration have used the nonparametric approach, specifically the Kaplan-Meier (KM) estimator of the survivor function (Besedeš and Prusa 2006a, Besedeš and Prusa 2006b, Nitsch 2009). The KM estimator is defined as:

$$\hat{S}(t) = \prod_{j \mid t_j < t} \left(\frac{n_j - d_j}{n_j} \right) \tag{1.3}$$

where nj is the number of individuals at risk at time tj and dj is the number of failures at time tj. The KM estimate is the product of all failure times less than or equal to t.

[14] For a detailed explanation of the parametric approach, see Cleves *et al.* (2010). This approach will not be explained here since its assumption of parameterization of both the hazard function as well as the impact of the covariates makes it of limited usefulness in our problem. Also, to the best of our knowledge, it has not been used in the trade duration literature.

The main limitation of the nonparametric approach is that only pair-wise comparisons can be considered. If there is heterogeneity among the individuals in each subgroup, it is not possible to hold any additional factors constant (e.g., *ceteris paribus*) when making the pair-wise comparison. It is also difficult to make comparisons across subgroups of factors that are continuous. For example, export duration may be affected by the value of exports, transportation costs, tariff rates, U.S. and exporting country GDPs, exporting country's share of production, U.S. price relative to exporting country price, exporting country price relative to the global price, and the U.S. consumption to production ratio – all of which are continuous variables. However, the KM estimates of the survivor function may still give some useful insights on how the survivor function differs between country-commodity pairs for binary variables of interest. As previously mentioned, SPS treatment requirements and NMA received during the study period are two of the binary variables of interest in our study. KM estimates may help us understand how the survivor function differs between country-commodity pairs that have SPS treatment requirements and those that do not and between country-commodity pairs that received NMA between 1996 and 2008 compared to those that have had market access throughout the sample period.

The main part of our analysis using the survival analysis method will be focused on the semiparametric approach. In semiparametric models (i.e. Cox PH model) the effect of the covariates is parameterized to alter the baseline hazard function (Cleves *et al*., 2010). However, the survivor function is not given a parametric form; hence we do not need to assume that the baseline hazard function has a specific shape (Royston and Lambert, 2011). This model therefore allows us to understand how covariates impact the hazard function. Suppose we only have one covariate x in our analysis, then the hazard function can be expressed as a product of two functions: $h(t, x, \beta) = h_0(t)r(x, \beta)$, where $h0(t)$ is the baseline hazard function which expresses how the hazard function changes as a function of survival time.[15] The function $r(x, \beta)$ expresses how the hazard function changes as the covariates change (Hosmer *et al.*, 1999). The hazard function may equal the baseline hazard function if $r(x = 0, \beta) = 1$, which in our case would mean that the value of all the covariates equals zero. We cannot imagine a case where variables such as GDP, price levels, production levels, etc., all equal zero. Hence, that case is not interesting to us, and it's not interesting in most empirical applications. Instead, the Cox PH model allows us to gain an understanding of the hazard function and the impact of the covariates without even estimating the baseline hazard function. Using the notation of Hosmer *et al.* (1999) consider the ratio of the hazard functions for two subjects 1 and 2 and one explanatory variable which takes the values x_1 for subject 1 and x_2 for subject 2:

$$HR(t, x_1, x_2) = \frac{h(t, x_1, \beta)}{h(t, x_2, \beta)}$$

$$HR(t, x_1, x_2) = \frac{h_0(t)r(x_1, \beta)}{h_0(t)r(x_2, \beta)}$$

$$HR(t, x_1, x_2) = \frac{r(x_1, \beta)}{r(x_2, \beta)} \tag{1.4}$$

Therefore, the hazard ratio depends only on the function $r(x, \beta)$. Given this insight, Cox (1972) proposed a model with the following parameterization of the impact of the covariates:

$$h(t, x, \beta) = h_0(t)e^{t\beta} \tag{1.5}$$

where the hazard ratio is:

$$HR(t, x_1, x_2) = e^{\beta(x_1 - x_2)} \tag{1.6}$$

Hence, when the explanatory variable x is a binary variable (for example, let x denote the existence of a free trade agreement (FTA) between the U.S. and the exporting country, hence

[15] The explanation can very easily extend for multiple covariates.

$x = 1$ means there is a FTA in place, and $x = 0$ means no FTA exists between the trade partners), then the hazard ratio estimated by the Cox PH model becomes: $HR = e^{\beta}$. If the value of the estimated coefficient is: $\beta = \ln(2)$, then the interpretation is that spells of service for country-commodity pairs for which there is a FTA in place are failing (ending) at twice the rate of those for which there is no FTA in place. Note that both the hazard function as well as the hazard ratios can be easily extended to include multiple covariates.

Let x denote the vector of explanatory variables included in our models. Following the literature on duration, we estimate a Cox PH model and consider various explanatory factors. We consider the importing and exporting countries' GDPs, transportation costs and tariff rates since such variables were found to impact trade duration in the literature (see for example Besedeš and Prusa 2006b). We also explore the impact of the average value of exports (see for example Besedes 2011a), as well as commodity prices in the U.S. and in the exporting country. Factors such as the average number of commodities a country exports to the U.S. (to measure experience with the specific partner) as well as number of exporting countries serving the U.S. market (to measure competition) were also found to be significant in the literature, and hence are included in our models (see for example Besedeš 2008 and Besedeš 2011a). In addition, we explore the impact of free trade agreements, SPS treatment requirements and NMA access on trade duration. We also try to capture whether exports from Mexico and Canada, which are both contiguous to the U.S. and also have a FTA in place with the U.S., have a different survivor experience compared to exports from other countries (several studies include regional binary variables as well as binary variables to account for common borders between the importing and exporting countries). Finally, we explore whether fruits and vegetables have different survival experiences by including a binary variable indicating whether the commodity is a fruit or a vegetable.

An issue that deserves some attention in the discussion of the Cox PH model, is that of data censorship and truncation. While there are several types of data censorship and truncation (Cleves *et al.*, 2010), only two types are applicable in our case: right censoring and left censoring. Right censoring occurs when a spell of service is not observed to have failed (ended) by the time the study period ends. Left censoring occurs when a spell begins before the analysis period (i.e. 1996 in our sample). While the Cox PH model can handle right-censored observations, it is less clear from the literature how to properly account for left-censored spells. One approach to account for left-censored spells is to estimate the hazard function using different variations of the sample (i.e. benchmark sample, gap-adjusted sample, first spell only sample, single spells only sample, etc.) in order to see if the results are robust across samples (Besedeš 2008, Nitsch 2009). A second approach is to collect data for the year prior to the year when the analysis period starts in order to identify which spells of service begin when the study period begins and which were already in existence when the study period begins. Then all observations with spells of service that start before the first year of the study period are dropped since those are left-censored observations (Besedeš, 2011a).

We will follow the first approach and estimate the Cox PH model using four variations of our sample: the full sample, first spell of service only, left-censored observations excluded, and gap adjusted sample that "fills" all one year gaps between two spells of service. In the gap adjusted sample, the one year gaps between two different spells of service are filled by assuming that exports took place during the gap year as well and that the one-year gap is an error (Besedeš and Prusa 2006b, Besedeš 2008). That way, if a spell of 2 years and a spell of 5 years are divided by a one year gap, in the gap adjusted sample, they are counted as only one spell which is 8 years long (2+1+5).

While the Cox proportional hazard model is useful because it allows us to understand the impact of the explanatory variables x, *ceteris paribus*, Hess and Persson (2010) identify several limitations, in its use for trade duration analysis. They recommend a discrete-time model specification based on the argument that discrete-time methods do not face the same limitations of continuous-time methods which were outlined in the Introduction section. Hence, in addition to the Cox proportional hazard model, we employ several variations of a popular class of count data models in a discrete time setting and compare the results between the two approaches. This approach is introduced next.

1.3.2. Count Data Model

Consider a random variable $Y(t)$ which describes the number of occurrences during the interval $(0, t)$, where $t > 0$. The count data framework is used to model $Y(T)$ for a specific time interval (T). For a fixed interval of time, if the probability distribution of Y(T) is a Poisson distribution, then a Poisson regression model is appropriate.[16] In cross sectional data, consider a sample of n observations, where the ith observation is denoted $(y_i, \boldsymbol{x}_i)$. The dependent variable y_i is a scalar variable denoting the number of occurrences of the event while $\boldsymbol{x}_i$ is a vector of independent variables that are thought to be determinants of the value of y_i. Then, y_i given $\boldsymbol{x}_i$ has a Poisson distribution with the following density function:

$$f(y_i|\boldsymbol{x}_i) = \frac{e^{-\mu_i}\mu_i^{y_i}}{y_i!} \tag{1.7}$$

where μ_i is determined by the set of independent variables in the Poisson regression model. For example, in the log-linear model specification μ is parameterized as follows:

[16] Winkelmann (1994) and Cameron and Trivedi (1998) offer an extensive review of Count Data Models.

$$\mu_i = \exp\,(\boldsymbol{x}_i'\boldsymbol{\beta}) \tag{1.8}$$

where one can easily derive the expected value of y_i based on the set of the independent variables as well as the set of estimated coefficients.

We estimate a Poisson regression model where the dependent variable is defined as the total number of years the U.S. imports a commodity from a given country. In the case of multiple spells, the number of years is added across all spells of service. For example, Table 1.3 shows that Argentina exported for 13 years, Australia for 3 years, Brazil for 8 years, and so on. Hence, the dependent variable does not include information on the number of spells of service. A similar set of explanatory variables used in the Cox PH model is also used in the Poisson model, following the literature on trade duration. More specifically, we estimate the following model:

$$E[y|\boldsymbol{x}] = exp(\beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k) \tag{1.9}$$

where x is a vector of explanatory variables specified in the previous section. Note that the coefficients of the independent variables are estimated by maximizing the respective log-likelihood function:

$$\ln L(\beta_i) = \sum_{i=1}^{n} [-\exp\,(\beta_i x_i + (\beta_i x_i) y_i)] \tag{1.10}$$

The main limitation of the Poisson distribution model is the independence assumption.[17] Namely the probability of an event happening in a certain time interval is independent of an

[17] Another important limitation in using the Poisson regression model is the assumption of the equality of the mean and variance, referred differently as the "equidispersion." However, Cameron and Trivedi (1998) argue that in real life data this assumption rarely holds. In most cases, they argue, there is overdispersion (underdispersion) where the variance is greater than (less than) the mean.

event happening in another non overlapping time interval. For example, if the variable of interest is the count of the number of articles a scholar has published, the assumption of independence implies that past success in publishing a paper does not affect the probability of future success (Long, 1997). In our case, this means that the event of a country exporting a commodity to the U.S. in 1996 is independent from the event of the same country exporting the same commodity to the U.S. in 1997. This is highly unlikely since the probabilities of these events occurring are not independent. In order to test this assumption, we will employ a probit model which seeks to explain what factors determine the probability that a country exports a given commodity in a given year. Duration (in years) of the relationship (up to that year) is included as one of the explanatory variables. We will also consider additional independent variables in the model specification, including factors of special interest in the case of FF&V imports, namely NMA and SPS treatment requirements.

1.3.3. Probit Model

The probit model is used to model binary response outcomes. Suppose one is interested to model the probability that an event occurs. A modeling technique that would only allow the probability to be strictly between 0 and 1 is necessary. While Linear Probability Models can be used to model binary dependent variables, they cannot assure that the predicted value of the dependent variable will be in the range $[0,1]$. Hence, probit and logit models are more appropriate to deal with such situations. Consider the following model:

$$P(y = 1|\boldsymbol{x}) = G(\beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k) = G(\beta_0 + \boldsymbol{x\beta}) \qquad (1.11)$$

Where the function G takes on values $0 < G(z) < 1$, for all real numbers z (Wooldridge, 2009).

The functional form of G in the logit model is logistic and in the probit model is normal. Hence, in the probit model, G is the standard normal cdf expressed as an integral.[18] The probit model can be derived from the latent (unobserved) variable model. Let y^* be a latent variable determined by:

$$y^* = \beta_0 + \boldsymbol{x\beta} + e, y = 1[y^* > 0], y = 0\ [y^* < 0] \tag{1.12}$$

Hence, the probability response P is defined to be between 0 and 1:

$$P = P(y = 1|\boldsymbol{x}) = P(y^* > 0|\boldsymbol{x}) = G(\beta_0 + \boldsymbol{x\beta}) \tag{1.13}$$

This provides the framework to conduct empirical estimations using the probit model. Note that since this is not a linear model, OLS estimation is not appropriate. Instead, maximum likelihood estimation (MLE) is used (Wooldridge, 2009).

We seek to understand the factors impacting the probability that a country will export a certain commodity to the United States. In order to answer this question we will make use of the panel dataset, which is an advantage of the probit model since we do not need to sum across years to get a count of years exporting. The analysis will still focus on the same commodities and exporting countries, in the period 1996-2008. The dependent variable (y) equals one if the export value of commodity i from exporting country j in year k was greater than \$10,000 (>\$0), and equals zero otherwise.[19] The \$10,000 threshold was used in the duration literature for robustness check of the result (Besedeš 2008, Nitsch 2009). We also use this threshold since small valued trade flows introduce noise to the empirical analysis and are often a result of contractual arrangements for research or specialty products.

[18] We concentrate the discussion on the probit model (rather than the logit) since this is the method we will use. According to Wooldridge (2009) this is the preferred method of most economists since the normal distribution assumption has some useful properties.

[19] Note that for robustness check we use two variations of the dependent variable: in the benchmark analysis we consider \$10,000 as the cut-off point (hence if the export value is below this threshold we code that as no exports); and in the alternative case we consider \$0 as the threshold value (so that any exporting value greater than zero indicates that exports did take place).

1.3.4. Data

Part of the data used in the empirical analysis of this study was compiled by Karov *et al*. (2009) and Jankovska *et al.* (2011) from a variety of sources outlined in these studies. It includes data on the value of exports per each country-commodity pair throughout the study period, tariff rates, transportation costs, SPS treatment requirements, NMA access cases issued during the study period, existence of FTAs between the U.S. and exporting countries, U.S. and exporting countries' GDPs, commodity prices in the U.S., in the exporting countries and the global averages, exporting countries' global share of production for each commodity, as well as characteristics of exporting countries such as contiguity to the U.S. and the region where they are located.

Based on this data as well as additional export and import data retrieved from the UN Comtrade database, we have generated several additional variables. Such variables include: price ratios for each commodity (between the U.S. and exporting country, and between the exporting country and the global price), U.S. consumption to production ratio for each commodity (where consumption equals: U.S. production + Imports – Exports), average number of exporters for each commodity, and average number of commodities exported to the U.S. by each exporting partner. Note that while for the Cox PH model and for the Poisson model the data for variables that vary over time are averaged out, the Probit model makes use of the panel dataset. For the Cox PH model, the variables are averaged across all the years for which a spell of service lasts. For the Poisson model, the data is averaged across all the years a trade relationship is in existence (hence, it excludes the gap years when the exports equal zero). This leads to a sample that varies between 693 and 1,000 observations for the Cox PH model, a sample that varies between 1,062 and 1,117 observations for the Poisson model, and a sample that varies between 4,594 and 8,454 observations for the Probit model. Note that sample size variations in each model are due to missing data for certain variables that are included in the different specifications (or, in the case of the Cox PH model, due to the variations in the sample itself).

1.4. Results

1.4.1. Survival Analysis

A nonparametric approach allows one to make pair-wise comparisons of the survival function between different groups within the sample. For example, does the survival function differ between exporter/commodity pairs subject to an SPS treatment (e.g., fumigation with methyl bromide or cold treatment) as a condition of entry into the U.S. compared to those that are not? As shown in Figure 1.1, for the 1,216 country/commodity/spell observations in our sample, the conditional probability of survival is higher for the country/commodity pairs with SPS treatment requirements (*Treat*) compared to those that do not. This may indicate that countries which already have invested in the necessary technologies to perform the required SPS treatments have an incentive to remain in the market for a longer duration. Note that the difference between the two groups is statistically significant using both the log-rank and Wilcoxon tests.

One of the provisions of the Agreement on Agriculture (AoA) from the Uruguay Round of WTO negotiations is that member countries provide increased market access to other member countries. Between 1996 and 2008, the U.S. has granted new market access to 67 country/commodity pairs (Jankovska *et al.*, 2011). Figure 1.2 shows how the survival function differs between country/commodity pairs that were granted new market access between 1996 and 2008 (*NMA*) and those country/commodity pairs that had market access throughout the period. The survival rate decreases more slowly for country/commodity pairs that gained new market access during the study period. However, the difference between the two groups is not statistically significant according to the log-rank hypothesis test, likely due to the low number of observations associated with *NMA*.

Figure 1. 1. Survival Function by *Treat*

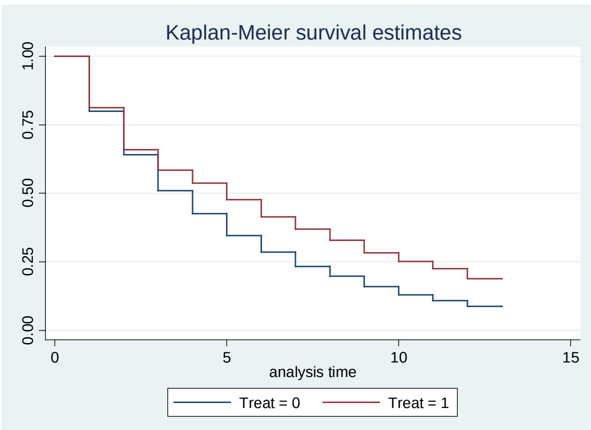
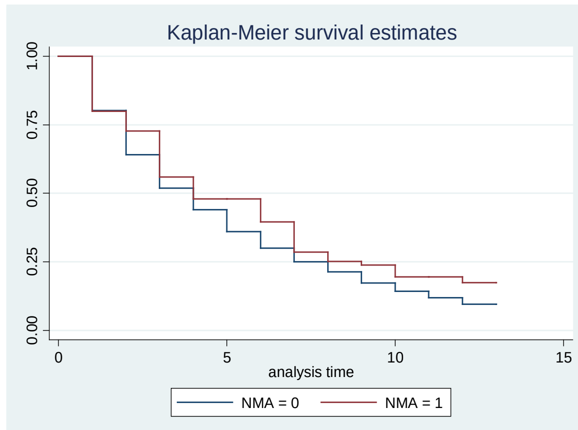


Figure 1. 2. Survival Function by *NMA*



As with any non-parametric analysis, the pair-wise comparisons are not able to hold all other variables constant. So while one can identify correlations between the survivor function and the variable of interest, it is not possible to prove causality. To do so, requires a semi-parametric method. We estimate a non-stratified Cox PH model, where a set of explanatory variables are used in order to explain the changes in the hazard rate.[20] Table 1.4 contains the results for each variation of the sample considered. Table B.1 in Appendix B.1 contains variable descriptions along with units of measurements.

Note that if a coefficient is positive (negative), the hazard rate increases (decreases) by the associated percentage.[21] Both binary as well as continuous explanatory variables are included in the model. The binary variables include: *Fruit, MexicoCanada, FTA, Treat,* and *NMA.* The continuous variables include: transportation costs, tariff rates, U.S. GDP, exporting country

[20] We also estimated stratified Cox PH models stratified by spell that yielded similar results.
[21] The percentage increase or decrease is calculated by taking the exponential of the estimated coefficient, both for binary variables as well as for continuous variables (Cleves *et al.*, 2010).

GDP, average value of imports, exporting country's global share of exports for the given commodity, price ratios between the U.S. and exporting country as well as between the exporting country and the world, U.S. consumption to production ratio, the number of countries exporting the given commodity to the U.S. (in order to measure competition), and the number of products that the exporting country exports to the U.S. (in order to measure reliability and experience with a partner).

Table 1. 4. Cox Proportional Hazards Model Results

	Full Sample	Gap Adjusted	First Spell Only	Left-Censored Dropped
	$\hat{\beta}$	$\hat{\beta}$	$\hat{\beta}$	$\hat{\beta}$
Fruit	0.088	0.054	0.123	0.144
	(0.125)	(0.144)	(0.155)	(0.142)
MexicoCanada	-0.376	-0.652	-0.583	-0.660
	(0.429)	(0.485)	(0.523)	(0.544)
FTA	0.412*	0.482**	0.113	0.389
	(0.243)	(0.233)	(0.347)	(0.256)
Treat	-0.184	-0.206	-0.110	-0.257
	(0.175)	(0.196)	(0.213)	(0.197)
NMA	-0.208	-0.380	0.235	-0.580*
	(0.279)	(0.319)	(0.270)	(0.319)
AvgValue	-0.002***	-0.002**	-0.002***	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)
AvgTC	-0.449*	-0.559**	-0.463	-0.304
	(0.261)	(0.222)	(0.357)	(0.332)
AvgTAR	0.606	0.221	1.650	0.515
	(0.916)	(0.903)	(1.320)	(1.014)
AvgUsGdp	-0.000***	-0.001***	-0.001***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
AvgExCoGdp	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
AvgExCoShare	1.193**	1.676***	0.911	1.537***
	(0.513)	(0.536)	(0.634)	(0.507)
AvgUSExCoPriceR	0.160**	0.195***	0.205***	0.141*
	(0.071)	(0.073)	(0.078)	(0.073)

Table 1.4. Continued

	Full Sample	Gap Adjusted	First Spell Only	Left-Censored Dropped
	$\hat{\beta}$	$\hat{\beta}$	$\hat{\beta}$	$\hat{\beta}$
AvgExCoWorldPriceR	0.000	0.000	0.000	0.000
	(0.000)	(0.001)	(0.001)	(0.000)
AvgUsConsProdR	0.092	-0.047	0.079	0.080
	(0.217)	(0.247)	(0.270)	(0.242)
AvgNoOfComm	-0.024*	-0.020	-0.013	-0.001
	(0.013)	(0.015)	(0.017)	(0.013)
AvgNoOfExp	-0.024*	-0.022	-0.029*	-0.018
	(0.013)	(0.014)	(0.016)	(0.015)
Observations	1,000	856	685	693
Log pseudolikelihood	-3,095.135	-2,371.658	-1,915.987	-2,310.436

Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1

Note: Left-Censored Dropped sample excludes all the spells that started in 1996, rather than only those that started before 1996. Only the cases where the value of imports exceeded $10,000 were included.

Across the different sample specifications, the impacts of variables whose estimated coefficients are statistically different than zero have the same direction.[22] The difference in the hazard rate for spells of service involving fruits versus vegetables is not statistically significant in any of the sample variations considered. Spells of exports from the neighboring countries of Mexico and Canada experience a lower hazard compared to other countries. In the benchmark sample, the hazard rate is 31% lower, although the result is not statistically significant. However, countries in a free trade agreement with the U.S. (excluding the case of NAFTA) experience a higher hazard rate. The results are consistent across sample variations and statistically significant in two out of the four cases. The benchmark sample results indicate that if the exporting country and the U.S. have a free trade agreement in place, the hazard rate increases by 51%. The existence of a free trade agreement may give an incentive to exporters to begin exporting to the United States. However, if the exporter is not reliable or lacks the necessary infrastructure to produce and export large quantities of the commodity at a competitive price, likely the spell of service will not last long. In fact the countries captured by this variable include mostly small countries (such as: Costa Rica, Dominican Republic, Guatemala, Honduras, Israel, Nicaragua, etc.), which for the reasons stated above might likely not have long spell durations despite the fact that a free trade agreement is meant to facilitate trade.

The two variables of special interest in this case – SPS treatment requirements and NMA obtained during the study period, do not have statistically significant impact on the hazard rate. The results in the benchmark sample indicate that both lower the hazard rate by 17% and 19% respectively. This may indicate that once a country has invested in human resources and infrastructure to fulfill the SPS requirements; they are more likely to continue exporting the

[22] Note that since the coefficients for several variables (such as: *Island*, *Landlocked*, *ColoLink*, *and CommOffLang*) were not statistically significant in other specifications of the Cox PH model considered, they were not included in the Cox PH model specification reported here.

given commodity. Similarly, if a country has completed the procedures necessary to receive the NMA (during the study period), they are likely to start exporting. However, the small number of cases where SPS treatments are required or NMA was issued during the study period, as well as the fact that the results are not statistically significant in the model suggest that these results are insufficient to derive a conclusion about the true impact of these factors.

As expected, spells of service involving a higher average value of exports have lower hazard rates. But contrary to what one would expect, higher transportation costs also lower the hazard rate. The impacts of both these variables are statistically significant across three and two sample variations respectively. The impact of an increase in the tariff rate however has the expected positive impact (leads to an increase in the hazard rate). As the GDP of the U.S. goes up, the hazard rate decreases. For example, an increase of $1billion in the GDP decreases the hazard rate by 0.05% according to the benchmark results. Albeit small, this relationship is expected since an increase in the GDP leads to a higher demand for all types of products, FF&V included. The GDP of the exporting country does not seem to have a significant impact on the hazard rate.

Exporting countries which have a larger share of exports of a commodity in the global exports of that commodity tend to have a shorter duration in the U.S. market. This result seems counterintuitive since the global share is an indication of reliability and experience. Another counterintuitive result is that of the ratio of the U.S. to the exporting country price. As the U.S. price goes up relative to that of the exporting country, the hazard rate increases. But one would expect that a higher price in the U.S. would make exports to the U.S. more attractive and hence lead to longer spells' duration. Both these results are statistically significant in three and four sample variations respectively. On the contrary, the results of the ratio of the exporting country price to the global price, as well as the consumption to production ratio in the U.S. are both not statistically significant.

Finally, better experience with the U.S. as measured by the number of commodities exported to the U.S., as well as more competition as measured by the number of exporting countries exporting a certain commodity to the U.S., lead to a lower hazard rate. The first result is expected because the more commodities an exporting country exports to the U.S., the better information partners in the U.S. have about exporters from that country hence the likelier they are to continue the relationship. However, more competition also decreases the hazard rate for all exporters already in the market. This may be as a result of a combination of issues. Certain exporters may have already built a brand name associated with their product (e.g. bananas from Panama) when other exporters enter the market, thereby differentiating their products. Furthermore, some consumers may have preferences on the geographic origin of the FF&V they purchase or preferences on the type of fruit or vegetable they purchase (e.g. the different varieties of apples), hence the sellers choose to provide several different options. For the exporters this in turn means that other exporters entering the market may not necessarily pose a big threat to their ability to survive in the market – although whether the value of exports would go down for all exporters as competition increases is a matter of further investigation.

Overall, the Cox PH model offers results that are not always intuitive. Furthermore, some of the variables of special interest in this case (*Treat* and *NMA*) were found to not be statistically significant. Although most of the results are robust across several variations of the sample, there are non-negligible limitations to using the Cox PH model which were discussed in more detail in the previous section. Hence, we conduct another robustness check by also estimating a Poisson regression (count data) model.

1.4.2. Count Data Model

The Poisson regression model is one of many count data models. As argued in the previous section, the dependent variable in the Poisson model is the number of years a country has exported a certain commodity to the U.S. during the study period.[23] This approach is different from the Cox PH model approach, where the dependent variable is a measure of the probability of failing. Hence, the dependent variable in the Poisson model is a discrete variable that takes on values from 0 to 13. A set of explanatory variables are used in three different model specifications. Specification 1 of the model only includes binary variables: *NMA*, *Treat*, *MexicoCanada*, *FTA*, and *Fruit*. Specification 2 includes the same explanatory variables that were used in the Cox PH model. Finally, several additional variables to identify specific characteristics of the exporting countries were added. Two of these additional variables identified geographic characteristics of the exporting countries (island and landlocked), and two variables were used to identify any cultural ties between the exporting country and the U.S. (historical colonial links and common official language). Since our main question of interest is whether or not the results from the Poisson model are similar to those obtained from the Cox PH model, we will focus on Specification 2 of the model. However, the results from Specification 1 and 3 are also reported for comparison.

[23] Similar to the Cox PH model, in this case as well, we consider \$10,000 of exports per year as the threshold value. Exports that have values less than \$10,000 are considered as \$0.

Table 1. 5. Count Model (Poisson) Results

	Specification 1	Specification 2	Specification 3
	$\hat{\beta}$	$\hat{\beta}$	$\hat{\beta}$
Fruit	0.842***	0.025	0.021
	(0.027)	(0.029)	(0.029)
MexicoCanada	1.710***	-0.014	-0.068
	(0.033)	(0.082)	(0.088)
FTA	1.116***	0.194***	0.182***
	(0.032)	(0.039)	(0.039)
Treat	0.530***	0.316***	0.316***
	(0.035)	(0.035)	(0.035)
NMA	0.160***	-0.172***	-0.165***
	(0.051)	(0.051)	(0.051)
AvgValue		0.000	0.000
		(0.000)	(0.000)
AvgTC		0.107**	0.095**
		(0.046)	(0.046)
AvgTAR		0.350**	0.281**
		(0.137)	(0.139)
AvgUsGdp		0.000	0.000**
		(0.000)	(0.000)
AvgExCoGdp		0.000	0.000
		(0.000)	(0.000)
AvgExCoShare		1.058***	1.190***
		(0.111)	(0.116)
AvgUsWorldPriceR		-0.004	0.001
		(0.017)	(0.017)
AvgUsConsProdR		0.000	0.000
		(0.000)	(0.000)
AvgNoOfComm		0.036***	0.037***
		(0.003)	(0.003)
AvgNoOfExp		0.050***	0.050***
		(0.004)	(0.004)
Island			0.100**
			(0.042)
ColoLink			-0.179***
			(0.064)
Landlocked			-0.518***
			(0.191)
CommOffLang			0.092**
			(0.036)
Observations	1,117	1,062	1,062
Log likelihood	-5,255.414	-3,519.519	-3,501.366

Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1

Exports of fruits do not have a statistically different duration compared to vegetables, and neither do exports from the neighboring countries of Mexico and Canada compared to the rest of the world. However, the impact of *Fruit* is positive while the impact of *MexicoCanada* is negative. These two factors were not found to be statistically significant in the Cox PH model either. The existence of a free trade agreement (FTA) between the U.S. and the exporting countries has a positive and statistically significant impact on the duration of the relationship.[24] If a *FTA* is in effect, the duration is 0.21 years longer than if a *FTA* is not if effect. By comparison, the existence of a *FTA* leads to a higher hazard rate in the Cox PH model – hence the results between the two models do not coincide in this case.

The two variables of special interest in our case: *Treat* and *NMA* – have opposite impacts on duration. Compared to the case when SPS treatments are not required, country-commodity pairs for which there are SPS treatment requirements experience a longer duration by 0.37 years. However, if the new market access (*NMA*) has been issued during the study period compared to before the study period, duration decreases by 0.16 years. Both these variables are statistically significant in the Poisson model, but where not statistically significant in the Cox PH model. The results for *Treat* coincide between the two models since the Cox PH model finds that SPS treatment requirements decrease the hazard rate. Conversely, the impact of *NMA* does not have the same direction on the two models considered.

To explore further the impact of NMA, we calculate the probability for each count of years/duration (ranging from 0 to 13) for countries issued new market access during the study period (NMA=1) versus those issued new market access before the study period (NMA=0). Holding all other factors at their mean values, the distributions are relatively similar in the two cases considered. If NMA=1, the count of years with the highest probability is 4 years (19.4%)

[24] This variable excludes NAFTA.

followed by 3 years (18.4%). If NMA=0, the count of years with the highest probability is 5 years (17.6%) followed closely by 4 years (17.5%). We also compare the impact of SPS treatment (*Treat*) in the group where the commodity exported is a fruit (*Fruit*=1), the country has a free trade agreement in effect (*FTA*=1) but is not Mexico or Canada (*MexicoCanada*=0) and the market access was granted before the study period (*NMA*=0). Holding all other factors at their mean values, for this group, if treatment is required (Treat=1) the count of years with the highest probability is 7 (14.4%). If treatment is not required (Treat=0) the count of years with the highest probability is 5 (16.9%).

The average value of exports does not have a significant impact on the duration of a trade relationship. Higher transportation costs and tariff rates, counter-intuitively both lead to an increase in duration. A one unit increase in the transportation costs leads to an increase of 0.11 years in duration, whereas a one unit increase in the tariff rate leads to an increase of 0.42 years in duration. By comparison, the Cox PH model finds that higher transportation costs decrease the hazard rate, but higher tariff rates increase the hazard rate. The U.S. GDP and the GDP of the exporting country do not have a statistically significant impact on duration according to the Poisson model results. A higher share of global exports by the exporting country leads to a higher duration in the trade relationship with the United States. A one unit increase in the global share of exports leads to an increase of 1.88 years in duration. However, a higher share leads to an increase in the hazard rate in the Cox PH model. The ratio of the U.S. price to the global price, and the U.S. consumption to production ratio have opposite impacts on duration, however neither one is statistically significant in the Poisson model. Finally, consistent with the findings from the Cox PH model, a higher number of commodities exported to the U.S. (measuring experience) as well as a higher number of exporters for a certain commodity (measuring competition) – both increase duration. Increasing the number of commodities exported to the U.S. by 1 additional commodity, increases duration by .04 years. Increasing the number of exporters of a certain commodity to the U.S. by an additional exporter leads to an increase of .05 years in duration. Similarly, the Cox PH model finds that both these factors decrease the hazard rate.

The results between the two models are not entirely consistent. Factors such as: SPS treatment, transportation costs, number of exporters, number of commodities, etc. – have consistent impacts on duration across the two models since the direction of the impact on duration is the same. Other factors, such as: attaining new market access during the study period, tariff rates, and exporting country global share of exports, etc. – have opposite impacts on duration across the two different models considered. However, some of these factors are not statistically significant in one of both of the models considered. For example, the coefficient on the tariff rate is statistically significant in the Poisson model but not statistically significant in the Cox PH model. On the other hand, the coefficient on the binary variable *Fruit* is not statistically significant in any of the two models considered, indicating that holding other factors constant, there is no difference in the survival experience between fruits and vegetables.

Since both models have several limitations as discussed in the previous section, further investigation of the impact of the factors that were found to have opposing impacts is needed. However, despite the limitations of the models, the impact of some of the factors is consistent across varying model specifications – hence we were still able to derive some important insight on trade duration in FF&V. In addition to understanding what factors impact trade duration, our goal is to also explore what factors impact the decision to enter the market. Are the same factors as significant in affecting the probability to export as they are in affecting the duration of the export relationship? In order to answer this question we estimate a Probit model the results of which are reported below.

1.4.3. Probit Model

The results from three different specifications of the Probit model are reported in Table 1.6. The dependent variable in the Probit model is a binary variable indicating whether or not a country has exported a given commodity to the U.S. in any given year during our study period.[25] We concentrate on Specification 3 since it includes all the variables that were explored in the previous models; however that comes with the expense of a smaller sample size as a result of missing data for certain variables. Note that for this model we make use of a Panel dataset.

[25] For the benchmark results reported, the threshold value is $10,000 per year; any values less than the threshold are considered $0. We also estimated the model where $0 is the threshold; the results do not vary significantly and are not reported here.

Table 1. 6. Probit Model Results

	Specification 1	Specification 2	Specification 3
	$\hat{\beta}$	$\hat{\beta}$	$\hat{\beta}$
Treat	0.154***	0.299***	0.353***
	(0.040)	(0.076)	(0.079)
NMA	-0.176	-0.290	-0.289
	(0.109)	(0.188)	(0.176)
NMATreat	0.122	0.353	
	(0.180)	(0.382)	
MexicoCanada	0.918***	0.695***	0.738***
	(0.046)	(0.090)	(0.090)
Fruit	0.248***	0.338***	0.383***
	(0.029)	(0.050)	(0.056)
FTA	0.096**	0.049	
	(0.049)	(0.095)	
UsGdp	-0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)
ExCoGdp	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)
Island		0.027	
		(0.088)	
CommOffLang		-0.177**	-0.076
		(0.069)	(0.064)
ColoLink		-0.026	
		(0.094)	
Landlocked		-0.680***	-0.850***
		(0.217)	(0.239)
TC		-0.172***	-0.147***
		(0.047)	(0.048)
TAR		0.546***	-0.007
		(0.152)	(0.168)
UsWorldPriceR			0.008
			(0.035)
ExCoShare			1.383***
			(0.264)
UsConsProdR			-0.000
			(0.000)
NoOfExp			0.041***
			(0.006)
Observations	8,454	5,116	4,594
Log likelihood	-5549.6750	-1650.2964	-1438.2333

Standard errors in parentheses: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

SPS treatment requirements have a positive impact on the decision to enter the market. Similarly, trade relationships from Mexico and Canada have a higher likelihood to come into effect, and so do relationships involving fruits. Since fruits are perennial commodities the incentive is bigger for suppliers to enter the market. A larger U.S. economy as measured by the GDP also has a positive impact on the probability to export to the U.S. Higher prices in the U.S. market compared to global prices leads to a higher probability to enter the U.S. market, however the impact is not statistically significant. Both a higher share of global exports by the exporting country and more competition as measured by the number of exporters of a certain commodity to the US, increase the probability of exporting. These factors are both statistically significant.

Conversely, there are also a number of factors that negatively impact the probability of exporting to the United States. Obtaining new market access during the study period decreases the probability of exporting to the United States. An increase in the GDP of the exporting country also has a very small but statistically significant negative impact. Both common official language and if the exporting country is landlocked have a negative impact on the probability of exporting, however the impact of the first is not statistically significant. As expected, higher transportation costs and higher tariff rates both lower the probability of exporting to the United States. Finally, a higher U.S. consumption to production ratio decreases the probability to export to the U.S., however this counterintuitive result is not statistically significant.

Overall the Probit model indicates that some of the factors considered have a similar impact direction on the probability to export as on the duration of exports. Such factors include: SPS treatment, new market access, competition as measured by the number of exporters, etc. On the other hand, some factors have opposite direction on the probability to export compared to duration. The most significant of which are: transportation costs and tariff rates.

1.5. Conclusions

Since current exporters face a significant amount of entry and exit from the market, it is useful to also understand the factors that impact duration, for both current and potential exporters. From the point of view of U.S. policymakers, understanding what factors lead to higher duration (and hence more stable trade relationships) is useful when designing policies to encourage longer trade duration for trade partnerships involving FF&V. This in turn will benefit consumers in terms of more reliable fresh fruit and vegetable product supply and increasing the number of competitors in the import market place which should lead to lower prices.

In seeking to understand the factors that impact trade duration in U.S. FF&V imports, we estimated two different models: Cox PH model and the Poisson (count data) model. While there are limitations to both the models considered, the empirical results suggest that factors such as SPS treatment requirements and experience with a partner (as measured by the total number of commodities exported to the U.S. by the exporting country) lead to a higher duration (Poisson model) and lower hazard rate (Cox PH model). However, there are also factors, such as attaining a new market access during the study period, that have opposite impact on duration in the two models considered (although the impact was only statistically significant in the Poisson model). We also estimated a Probit model to understand what factors impact the probability of trading, and found that while some factors impact both the probability to trade as well as duration in the same direction (Treat, NMA, no of exporters, etc.), the impact of other factors (transportation costs and tariff rates) has an opposite direction.

The results indicate that for exporters that are interested in entering the U.S. market, finding a way to lower the transportation costs would lead to a higher probability of exporting. As the exporting country's share of global production increases, the exporting country is more likely to export that commodity to the United States. This indicates that exporting countries can increase the odds of exporting to the U.S. as they increase their global share of production for the specific commodity. However the impact of this variable on duration has opposite signs on the two models considered, hence our results are not sufficient to offer a definitive conclusion on how this factor impacts duration. It is also difficult to derive a definitive conclusion on the impact of FTAs on duration. The results from the Cox PH model indicate that the existence of a FTA between the U.S. and the exporting country increases the hazard rate, while the Poisson model results indicate a FTA leads to a larger number of years in the market. In both cases the results are statistically significant. Hence, further research is needed to understand the impact of this factor on duration. The probit model results indicate that the impact of FTA on the probability to export is positive. Hence while we cannot derive a definitive conclusion on how FTA impacts overall duration, countries that have a FTA with the U.S. have a higher probability to export.

Country characteristics (of exporting countries) such as: island, landlocked, colonial links with the U.S. and common official language with the U.S., were excluded from the Cox PH models since they were not statistically significant in explaining duration in FF&V. However, the Poisson model finds these factors to be statistically significant. For example, exporting countries that are islands have a higher number of years in the U.S. market, while being a landlocked country decreases the duration of the relationship. If the exporting country and the U.S. have any past or present colonial links, the duration of the relationship is shorter. However, a common language between the U.S. and the exporting country leads to a longer duration, since it possibly indicates cultural similarities and hence taste similarities between the two countries.

With the increase in the standard of living (among other factors), U.S. consumers are demanding an increasingly diverse and wide set of products. However, in the current age with extensive information readily available, consumers are increasingly concerned with the origin of the products they use and especially of the food they consume. Currently FF&V are supplied from various sources as exporting countries enter and exit the market with high frequency, hence possibly affecting the welfare of the consumers. While this study focuses primarily on the impact of the supply side factors, it would be interesting to explore how factors from the demand side affect trade duration. Such factors could try to capture consumer preferences, the role of information that consumers posses, etc.

Trade duration for foreign exporters may also impact the welfare of U.S. producers. With a high frequency of entry and exit into the market for foreign competitors, it is important for U.S. producers to understand the factors that impact trade duration. This would allow them to predict better the quantities of FF&V that need to be produced for the market at home.

Since both the Cox PH and the Count data models have limitations in the trade duration literature, future studies should explore whether other methodologies are more appropriate to be used in such applications. Since the impact of some variables was found to not coincide between the two models, future studies should also explore a longer study period which could provide more insight on the factors with mixed results. The application can also be extended to other types of commodities, such as meat products and dairy products. Finally, it would be interesting to estimate a Heckman two-step procedure that simultaneously estimates how covariates impact the probability of exporting as well as the intensity of exports. This would allow us to understand how duration impacts both these factors. However, to do so, an exclusion factors is necessary for the first stage equation. Such an exclusion factor would need to not affect the intensity of trade, but only the probability to trade. Future work should concentrate on identifying such a factor and carrying on the analysis.

References

Aksoy, M.A., and J.C. Beghin. 2005. "Global Agricultural Trade and Developing Countries." Washington DC: The World Bank.

Besedeš, T. (2008). "A Search Cost Perspective on Formation and Duration of Trade." *Review of International Economics* **16**: 835–849. doi: 10.1111/j.1467-9396.2008.00752.x

Besedeš, T. (2011a). "Export differentiation in transition economies." *Economic Systems* **35**(1): 25-44.

Besedeš, T. (2011b). "The role of NAFTA and returns to scale in export duration." MPRA paper number: 31304.

Besedeš, T. and T. J. Prusa (2006a). "Ins, outs, and the duration of trade." *Canadian Journal of Economics/Revue canadienne d'économique* **39**(1): 266-295.

Besedeš, T. and T. J. Prusa (2006b). "Product differentiation and duration of US import trade." *Journal of International Economics* **70**(2): 339-358.

Besedeš, T. and T. J. Prusa (2011). "The role of extensive and intensive margins and export growth." *Journal of Development Economics* **96**(2): 371-379.

Brenton, P., M. D. Pierola and E. von Uexküll (2009). "The Life and death of trade flows: understanding the survival rates of developing-country exporters", in R. Newfarmer, W. Shaw and P. Walkenhorst (eds.), *Breaking Into New Markets: Emerging Lessons for Export Diversification*, pp. 127–144, Washington D.C.: The World Bank.

Brenton, P., C. Saborowski, and E. von Uexkull (2010). "What Explains the Low Survival Rate of Developing Country Export Flows?" *The World Bank Economic Review* **24**(3): 474-499.

Cadot, O., L. Iacovone, F. Rauch, and M. D. Pierola (2011), "Success and Failure of African Exporters," World Bank Policy Research Working Paper Series.

Cameron, A. C. and P. K. Trivedi (1998). "Regression analysis of count data." Cambridge University Press.

Clemens, R. 2004. "The Expanding U.S. Market for Fresh Produce." Iowa Ag Review, Center for Agricultural and Rural Development. 10(1): 8-10.

Cleves, M., R. G. Gutierrez, W. Gould, and Y. V. Marchenko (2010). "An Introduction to Survival Analysis Using Stata." 3rd edition. College Station: A Stata Press Publication.

Cox, D. R. (1972). "Regression Models and Life-Tables." *Journal of the Royal Statistical Society. Series B (Methodological)* **34**(2): 187-220.

Hess, W. and M. Persson (2010). “The Duration of Trade Revisited: Continuous-Time vs. Discrete-Time Hazards.” IFN Working Paper No. 829.

Hess, W. and M. Persson (2011). "Exploring the duration of EU imports." *Review of World Economics* **147**(4): 665-692.

Hosmer, D. W. and S. Lemeshow (1999). “Applied Survival Analysis : Regression Modeling of Time to Event Data.” Wiley.

Jankovska, O., E.B. Peterson, J.H. Grant, G.C. Ferreira. 2011. “New Market Access in Fresh Fruits and Vegetable Imports to the United States.” Master’s Thesis, Virginia Polytechnic Institute and State University.

Jenkins, S. P. “Lesson 4. Estimation of the (integrated) hazard and survivor functions: Kaplan-Meier product-limit and lifetable methods.” <http://www.iser.essex.ac.uk/files/teaching/stephenj/ec968/pdfs/ec968st4.pdf>.

Johnson, R. 2010. "The U.S. Trade Situation for Fruit and Vegetable Products." Washington DC: Congressional Research Service Report for Congress.

Karov, V., E.B. Peterson, J.H. Grant, B. Mills, and J. Alwang. 2009. "Trade Barriers or Trade Catalysts? The Effects of Phytosanitary Measures on U.S. Fresh Fruit and Vegetable Imports." Master's Thesis, Virginia Polytechnic Institute and State University.

Long, J.S. (1997) “Regression Models for Categorical and Limited Dependent Variables.” Thousand Oaks: SAGE Publications.

Long, J. S. and J. Freese (2006) “Regression Models for Categorical Dependent Variables Using Stata.” 2nd edition. College Station: A Stata Press Publication.

Lucier, G., S. Pollack, and M. Ali. 2006. "Fruit and Vegetable Backgrounder." Washington D.C: Economic Research Service (ERS), U.S. Department of Agriculture (USDA).

Mills, B. F. (2000). "Are Spells of Unemployment Longer in Nonmetropolitan Areas? Nonparametric and Semiparametric Evidence." *Journal of Regional Science* **40**(4): 697-718.

Nitsch, V. (2009). "Die another day: duration in German import trade." *Review of World Economics* **145**(1): 133-154.

Obashi, A. (2010). "Stability of production networks in East Asia: Duration and survival of trade." *Japan and the World Economy* **22**(1): 21-30.

Royston, P., and P. C. Lambert. (2011). “Flexible Parametric Survival Analysis Using Stata: Beyond the Cox Model.” 1st edition. College Station: A Stata Press Publication.

Stata. FAQs. Analysis of multiple failure-time survival data. “How do I analyze multiple failure-time data using Stata?” <http://www.stata.com/support/faqs/stat/stmfail.html>.

Stata Topics: Regression with Count Data. Retrieved 2012 from http://www.ats.ucla.edu/stat/stata/topics/CountModels.htm.

United Nations Commodity Trade Statistics Database (UN Comtrade). 2010. Retrieved 2009 from http://comtrade.un.org/db/dqBasicQuery.aspx.

Winkelmann, R. (2000). “Econometric Analysis of Count Data.” 3rd edition. Springer.

Wooldridge, J.M. (2009) "Introductory Econometrics: A Modern Approach." Mason, OH: South-Western Cengage Learning.

Appendix 1.A.

Table A. 1. List of Exporting Countries of FF&V

#	Country	ISO3	#	Country	ISO3	#	Country	ISO3
1	Afghanistan	AFG	31	Ghana	GHA	61	Peru	PER
2	Argentina	ARG	32	Greece	GRC	62	Philippines	PHL
3	Australia	AUS	33	Grenada Is	GRD	63	Poland	POL
4	Bahamas	BHS	34	Guatemala	GTM	64	Portugal	PRT
5	Bangladesh	BGD	35	Haiti	HTI	65	Romania	ROU
6	Belgium	BEL	36	Honduras	HND	66	Russia	RUS
7	Belize	BLZ	37	Hong Kong	HKG	67	Saudi Arabia	SAU
8	Bolivia	BOL	38	Hungary	HUN	68	Serbia/Montenegro	SCG
9	Bosnia-Hercegov.	BIH	39	India	IND	69	Singapore	SGP
10	Brazil	BRA	40	Indonesia	IDN	70	South Africa	ZAF
11	Bulgaria	BGR	41	Iran	IRN	71	Spain	ESP
12	Cambodia	KHM	42	Ireland	IRL	72	Sri Lanka	LKA
13	Cameroon	CMR	43	Israel	ISR	73	St Lucia Is	LCA
14	Canada	CAN	44	Italy	ITA	74	St Vinc. & Gren.	VCT
15	Chile	CHL	45	Jamaica	JAM	75	Sweden	SWE
16	China	CHN	46	Japan	JPN	76	Switzerland	CHE
17	Colombia	COL	47	Korea	KOR	77	Syria	SYR
18	Costa Rica	CRI	48	Lebanon	LBN	78	Taiwan	TWN
19	Cote d`Ivoire	CIV	49	Macedonia	MKD	79	Tanzania	TZA
20	Croatia	HRV	50	Madagascar	MDG	80	Thailand	THA
21	Denmark	DNK	51	Malaysia	MYS	81	Tonga	TON
22	Dominican Rep.	DOM	52	Mexico	MEX	82	Trin. & Tobago	TTO
23	Ecuador	ECU	53	Morocco	MAR	83	Turkey	TUR
24	Egypt	EGY	54	Mozambique	MOZ	84	United Arab Em.	ARE
25	El Salvador	SLV	55	Netherlands	NLD	85	United Kingdom	GBR
26	Estonia	EST	56	New Zealand	NZL	86	Uruguay	URY
27	Ethiopia	ETH	57	Nicaragua	NIC	87	Venezuela	VEN
28	Fiji	FJI	58	Nigeria	NGA	88	Vietnam	VNM
29	France	FRA	59	Pakistan	PAK	89	Zimbabwe	ZWE
30	Germany	DEU	60	Panama	PAN			

(Karov *et al*., 2009; Jankovska *et al.*, 2011)

Appendix 1.B.

Table B. 1. Explanatory Variables' Description

Variable	Description
Treat	Binary; SPS treatment(s) are requirements.
NMA	Binary; New market access is issued during the study period.
NMATreat	Binary; Interaction term between *Treat* and *NMA*.
MexicoCanada	Binary; Exporting country is Mexico or Canada.
Fruit	Binary; Commodity is a fruit.
FTA	Binary; Free Trade Agreement between the US and Exporting Country is effect (NAFTA is excluded).
UsGdp	Continuous; U.S. annual GDP.
ExCoGdp	Continuous; Exporting Country annual GDP.
Island	Binary; Exporting country is an island.
CommOffLang	Binary; Exporting country and U.S. share an official language.
ColoLink	Binary; Exporting country and U.S. have or have had colonial ties.
Landlocked	Binary; Exporting country is landlocked.
Value	Continuous; Value of exports per year - Customs Value.
TC	Continuous; Transportation costs measured by (Cost, Insurance and Freight / Customs Value).
TAR	Continuous; Tariff Rate measured by (Landed Duty Paid / Cost, Insurance and Freight).

Table B.1 Continued

Variable	Description
UsExCoPriceR	Continuous; U.S. price to exporting country price ratio.
UsWorldPriceR	Continuous; U.S. price to world price ratio.
ExCoWorldPriceR	Continuous; Exporting country to world price ratio.
ExCoShare	Continuous; Exporting country share of world exports.
UsConsProdR	Continuous; U.S. consumption to production ratio. Consumption is ((US prod + imports)-exports).
NoOfExp	Discrete (Continuous when averaged out); Number of exporters for each commodity.
NoOfComm	Discrete (Continuous when averaged out); Number of commodities each partner exports.

Chapter 2 - Goods-Time Elasticity of Substitution in Household Food Production for SNAP participants and nonparticipants

2.1. Introduction

The Supplemental Nutrition Assistance Program (SNAP), previously known as the Food Stamp Program (FSP), was initiated in 1939 and has since seen fluctuation in the number of participants – however, it remains one of the key federal programs to alleviate hunger and malnutrition. In 2011, approximately 44.7 million Americans nationwide received food assistance through SNAP. The program is predominantly financed by the federal government, with state governments financing part of the administrative costs. That same year, the federal government's expenditures on SNAP reached $78 billion, 92 percent of which was spent on the benefits, while the remaining 8 percent was spent on administrative expenses (Center on Budget and Policy Priorities, 2012).

SNAP eligibility is determined through a set of three factors: income, assets, and household size and composition (United States Department of Agriculture – Food and Nutrition Service, 2012). The benefits from the SNAP program can be used to purchase food items as well as seeds and plants that lead to the production of food.[26] The benefits cannot be used to purchase alcoholic drinks and tobacco products, non-food items, vitamins and medicines and pre-prepared hot meals and foods that can be eaten in the store (USDA – FNS, 2012). With few exceptions, SNAP benefits cannot be used in restaurants and fast food chains (Laskawy, 2011).[27] Hence,

[26] The level of SNAP benefits is determined through the Thrifty Food Plan (TFP), which "provides a representative healthful and minimal cost meal plan" assuming that all meals are prepared at home (Carlson et al. 2007).
[27] Several states allow elderly, homeless and disabled persons receiving SNAP benefits to use such benefits in certain fast food restaurants.

participating households that depend mainly on SNAP benefits to purchase their food may need to spend more time preparing meals at home compared to non-participating households that do not face such restrictions.

The process of preparing a meal, whether it is a sandwich or a three course dinner, requires the necessary ingredients as well as time. In a typical household, both these inputs are limited, and in all households they come at a cost. As a result, a question of interest for public policies, and also for household technology analysts, is to what extent can food expenses be substituted with more time and vice versa in the process of home food production? Since a large number of households receive SNAP benefits and face the restrictions outlined above, a related question is whether (and if so to what degree) the goods-time elasticity of substitution in household food production differs between SNAP participants and nonparticipants? The answer to this question has important implications for policies related to SNAP. If the goods-time elasticity of substitution is low for SNAP participants, then just increasing the amount of benefits paid to households (as was done recently with the 2009 American Recovery & Reinvestment Act – ARRA provisions for SNAP) will not significantly increase food production at home, and hence food consumption and calorie intake if time that can be dedicated to food production remains limited. Hence, the degree of substitution for SNAP participants has important implications for the effectiveness of policies related to SNAP.

2.1.1. Objectives

The main objective of this study is to estimate the goods-time elasticity of substitution in household food production for SNAP participants versus nonparticipants. Following the literature, we will investigate whether the results vary and in what direction if time in food consumption is included ("eating") versus if time in food consumption is excluded from the analysis ("food production"). The policy implications of the results will also be discussed. We will use the concept of Morishima elasticity of substitution in order to estimate the goods-time elasticity. The dataset to be used for this analysis combines data from the following surveys: Current Population Survey (CPS), Food Security Supplement (FSP), and the American Time Use Survey (ATUS).

2.1.2. Organization

The study is organized as follows. Section 2.2 offers a summary of the literature both on goods-time elasticity of substitution in household food production, as well as on the SNAP program and program participation. Section 2.3 includes a discussion of the conceptual and the empirical framework of the study. A summary of the data and data sources is provided in section 2.4. The results of this study are provided in section 2.5 followed by a discussion on the study's conclusions, limitations and recommendations in section 2.6.

2.2. Literature Review

2.2.1. Goods-Time Elasticity of Substitution

Substitutability between food expenditures and time can be analyzed from several different perspectives depending on the motivation. In terms of health and nutrition as well as policies related to promoting healthy lifestyles and fighting various phenomena possibly related to food consumption (such as high levels of obesity in the population), the question of interest is to what extent is "eating" time (which includes both food consumption as well as food production) substitutable with more food expenses, ceteris paribus? Using household data from the American Time Use Survey (ATUS) matched with Current Population Survey (CPS) and the Food Security Supplement (CPS-FSS), Hamermesh (2008) estimates the goods-time elasticity of substitution in eating to be approximately 0.33 for the reference week food expenditures (ATUS diary day) and approximately 0.22 for the usual weekly food expenditures. These results show a relatively low level of substitution between "eating" time and food expenditures for the sample space which included married couples between 18 and 64 years old (Hamermesh 2008).

Conversely, household technology researchers are mainly concerned with the process of food production (e.g. how to increase home food production or its efficiency); hence, the inclusion of consumption time in the analysis can distort the simple production relationship. Baral, Davis and You (2011) show analytically that when consumption time is *not* included in the analysis (i.e., pure "food production" time) the goods-time elasticity of substitution is higher. Baral, Davis and You (2011) using the same data as Hamermesh (2008), find that the goods-time elasticity of substitution is about 60% greater when consumption time is not included. In essence, Baral, Davis and You (2011) show analytically and empirically a special case of what Diewert (1974) had shown analytically, that as inputs are disaggregated, given a set of assumptions, the expectation is to encounter a higher aggregate elasticity of substitution. Neither Hamermesh (2008) nor Baral, Davis and You (2011) estimate the elasticity of substitution for SNAP participants. However, as indicated in the previous section, understanding the goods-time elasticity of substitution for SNAP participants has important policy implications for policies related to SNAP. Hence, our main contribution to the literature is providing estimates of the elasticity of substitution (both when consumption time is included as well as when consumption time is excluded) for two different population subgroups: SNAP participants and nonparticipants.

2.2.2. Supplemental Nutrition Assistance Program Participation

In order to understand how the goods-time elasticity of substitution varies between SNAP participants and nonparticipants, we need to adjust for selection bias in observational data.

Selection bias arises as a result of the fact that SNAP participation is nonrandom in the population. As will be explained in-depth in the next section, we will make use of the propensity score matching methodology to estimate the treatment effect (SNAP participation). This will allow us to get consistent elasticity estimates, by controlling for factors that impact SNAP participation. As discussed below, factors that impact SNAP participation have been studied extensively in the past two decades.

Since its implementation started, SNAP has evolved in a lot of aspects: eligibility requirements, level of participation, level of coverage, level of benefits, the process of redeeming benefits, etc. A large portion of the population currently participates in the program. For example, in 2011, approximately 14% of the population received SNAP benefits at some point during the year. That same year, more than 21 million households participated in the program per month and received on average $284 in benefits per household per month (USDA-FNS, 2012). As will be explained in more depth in the Data section, our analysis period extends from 2006 to 2008. During that time period, the number of SNAP participants nationwide ranged from 26.3 million (2007) to 28.2 million (2008). The national average monthly benefits per participant during this period were: $94.75 (2006), $96.18 (2007) and $102.19 (2008) (USDA-FNS, 2012). As expected, in times of nationwide economic downturn the number of participants rises significantly (as in the last few years since 2008), while in times of economic expansion the number of participants declines (such as during mid- to late- 90s) (Kabbani and Wilde, 2003). However, the fluctuations in the level of participation cannot be explained by the performance of the economy alone. The literature (Currie and Grogger, 2001; Kabbani and Wilde, 2003; Mykerezi and Mills, 2010) on SNAP participation finds that the factors that impact participation levels can be categorized in three main categories: factors related to the performance of the overall economy (including: GDP growth, state unemployment rates, etc.), SNAP program design and administrative characteristics (recertification periods, the overpayment and underpayment benefits' levels, changes in eligibility requirements, transaction costs, etc.), and households' structures and characteristics.

Currie and Grogger (2001) considered how factors related to the performance of the economy as well as policies related specifically to FSP (now SNAP) affect participation. They found that changes in unemployment rates accounted for only about 20% in the changes on program participation during the study period considered, 1993-1998. These authors further argue that another assistance program – Temporary Assistance for Needy Families (TANF) – had a significant role in explaining the decline in FSP participation during that same period of time. The impact of TANF was higher for single headed households, a group which was also affected to a great extent from the short recertification periods (Currie and Grogger, 2001). Short recertification periods are implemented in cases where states seek to lower their respective error rates (underpayment and overpayment of benefits) (Kabbani and Wilde, 2003). Kabbani and Wilde (2003) found that such measures did in fact lead to lower error rates during the period of their analysis, 1990-2000. However, they also found that shorter recertification periods lead to lower FSP participation rates because of two main reasons: lower error rates lead to the elimination of non-eligible households, and shorter recertification periods increase the transaction costs of participation for eligible households (Kabbani and Wilde, 2003). Mykerezi and Mills (2010) find that a higher caseload overpaid leads to higher FSP participation, while a higher caseload underpaid leads to a lower FSP participation. They also find that higher state unemployment rates positively impact FSP participation (Mykerezi and Mills, 2010). However, several studies find that these factors do not have the same impact on all types of households – in fact different factors may encourage or discourage participation of different population subgroups in different levels (Currie and Grogger 2001, Kornfeld 2002, Kabbani and Wilde 2003). For example, Currie and Grogger (2001) find that the decline in the unemployment rate during the period 1993-1998, explained 16% of the fall in FSP participation for single mothers; but the impact was different for other subgroups of the population: married heads with children (26%), married heads without children (43%), households without children and with elderly members (13%) and adults living alone (44%).

Household composition and characteristics also impact SNAP/FSP participation. As expected based on eligibility requirements, household wealth and income are negatively related with the probability of participation, while months unemployed is positively related to the probability of participation (Mykerezi and Mills, 2010). Larger households, households headed by a female, and non-white households are more likely to be SNAP participants. Conversely, the education level and age of the household head are negatively related with program participation (Mykerezi and Mills, 2010).

The main goal of this study is to understand the impact of SNAP participation on the goods-time elasticity of substitution in household food production. In order to estimate the goods-time elasticity of substitution for the SNAP participants and nonparticipants, we will follow the analytical methodology used in the literature (Hamermesh 2008, Baral, Davis and You 2011) as explained below.

2.3. Conceptual and Empirical Framework

Following Hamermesh (2008) and Baral, Davis and You (2011), the conceptual framework of this study is based on the household production theory. The cost minimization problem can be expressed in general notations as: Minimize $C = \sum_{i=1}^{i=n} \boldsymbol{w} x_i$, subject to: $f(x_1 \dots x_n) = y'$, where y' is a predetermined level of food output to be produced, x_i denotes quantities of inputs and **w** is a vector of input prices (Silberberg 1990). The resulting cost function is a function of the total output level as well as the prices of the inputs used,

$$C = C^*(y, \boldsymbol{w}) \tag{2.1}$$

Note that in our case, food ingredients as well as time dedicated to the process of food production (and consumption) are considered as input factors and the price of time is considered the wage rate.

According to Blackorby and Russell (1981), the Morishima elasticity of substitution (MES) between inputs x_i and x_j can be expressed as follows:

$$MES_{ij}(y, w) = \frac{-\partial \ln \left(\frac{C_i(y, \boldsymbol{w})}{C_j(y, \boldsymbol{w})} \right)}{\partial \ln \ (w_i / w_j)} \tag{2.2}$$

$$= \frac{\partial \ln C_j \ (y, \boldsymbol{w})}{\partial \ln w_i} - \frac{\partial \ln C_i \ (y, \boldsymbol{w})}{\partial \ln w_i} \tag{2.3}$$

$$= \frac{\partial \ln x_j \ (y, \boldsymbol{w})}{\partial \ln w_i} - \frac{\partial \ln x_i \ (y, \boldsymbol{w})}{\partial \ln w_i} \tag{2.4}$$

Note that the subscripts of C indicate partial derivatives with respect to w_i or w_j. Going from equation (2.2) to (2.3), we use logarithmic rules, and going from equation (2.3) to (2.4) we use the Shephard's Lemma, namely: $x_i^H = C_i$. This is a useful result because it allows us to estimate the MES by estimating the input demand functions.

A general requirement of estimating demand functions is the need for a measure of output. In our case, the output is not specified in our data (i.e. the respondents are not asked about what type of food they prepared when they report to have engaged in food preparation).

However, if the underlying technology is homothetic the ratio of input demands and hence the log difference in inputs is independent of output.[28] This can be seen as follows. Homotheticity implies the cost function has the structure (Silberberg, 1990)

$$C = g(y)f(\boldsymbol{w}) \tag{2.5}$$

and so by Shephard's Lemma, we have:

$$x_i = g(y)\frac{\partial f(w)}{\partial w_i} = g(y)h_i(w) \tag{2.6}$$

where $h_i = \frac{\partial f(w)}{\partial w_i}$. Taking the logarithm and using (2.4), yields:

$$MES_{ij} = \frac{\partial \ln h_j(w)}{\partial \ln w_i} - \frac{\partial \ln h_i(w)}{\partial \ln w_i} \tag{2.7}$$

So in the homothetic case the *MES* is not a function of output. See also Davis and Shumway (1996).

Similar to Hamermesh (2008) and Baral, Davis and You (2011), we assume that the home food production function has a CES form. Following the empirical models of Hamermesh (2008) and Baral, Davis and You's (2011) and modifying appropriately for the hypothesis at hand, we will estimate the following:

$$\ln\left(\frac{x_f}{t_f}\right) = \sigma_0 + \sigma_1 \ln wage + \sigma_2 SNAP * \ln wage + \sigma_3 SNAP + \sigma_4 OwnChild + \varepsilon \tag{2.8}$$

where x_f denotes food expenditures while t_f denotes time spent preparing the meal.[29] Therefore, the estimated MES for SNAP nonparticipants will be $\hat{\sigma}_1$ while for SNAP participants will be $\hat{\sigma}_1 + \hat{\sigma}_2$.

[28] Refer to Silberberg (1990) for a review on the properties of homothetic functions.

Since the data on the wage variable is missing for approximately 82 percent of the sample, we employ a Heckman two-step procedure to impute the wage rate as in Baral, Davis and You (2011). In equation (2.8), $SNAP$ denotes a binary variable indicating whether the household has reported to have received SNAP benefits in the past 12 months or not.[30] However, SNAP participation is affected by a set of many factors, as discussed in the literature review section, some of which (such as: gender, employment status, income level, etc.) may also impact the dependent variable in (2.8). One way to adjust for selection bias when using observational data is the propensity score matching methodology (Wooldridge, 2010). This is a methodology often used in the literature to estimate the treatment effect when experimental data is not available (Abadie and Imbens, 2011).

In this study we implement the propensity score matching to estimate the treatment effect, where treatment is SNAP participation. Based on the findings in the literature discussed above, we include a set of covariates that were found to impact SNAP participation. State unemployment rates are included to account for the performance of the economy, while SNAP overpayment and underpayment rates are included as two important characteristics of the program. Additionally, household characteristics such as: level of education, employment status, age, gender, race, metropolitan status, number of children under 18 years old, and household

[29] We also estimate the MES for the "eating" occasion. In that case, consumption time is also included in the analysis.

[30] In an alternative specification of the sample, we separated non-eligible nonparticipants from eligible nonparticipants based on the survey universe of respondents for SNAP participation question. However, the universe of respondents for this question is only determined by the income factor (includes only respondents below 185% of the poverty line) and a question on food security (those that answer "Yes" to the question "In the 12 months, since December of last year, did you ever run short of money and try to make your food or your food money go further?"). Hence it disregards other factors that are taken into consideration when determining SNAP eligibility. The data does not otherwise give an indication on whether the household is eligible to receive SNAP benefits or not. When we reduce the sample to the eligible participants and nonparticipants based on the universe of this question, the sample size reduces to 915 observations however the results tell the same story as those obtained from the full sample.

income, are also included in the analysis. In order to carry on the empirical estimations we make use of data from a combination of surveys as explained in the next section.

2.4. Data

The empirical analysis outlined above requires data on the following variables: Time in household food production (time preparing the meal, time cleaning-up, time grocery shopping and time traveling for grocery shopping), as well as time spent in food consumption. Furthermore, data is needed on: the wage rate, food expenditures, SNAP participation status indicator, number of children below 18 years old, gender, race, age, level of education, household income, metropolitan status, and employment status. Additionally, state level data (including: state unemployment rates, SNAP over-payment rates and SNAP under-payment rates) are also necessary in order to carry on the empirical analysis. The data is obtained from three inter-related surveys as well as several other sources described below.

In accordance with previous studies on this topic, we will primarily employ data from the American Time Use Survey (ATUS) as well as the Current Population Survey – Food Security Supplement (CPS - FSS) in order to carry out the empirical study. ATUS data was retrieved from Bureau of Labor Statistics – ATUS (2008).[31] CPS-FSS data was retrieved from United States Census Bureau – CPS DataFerret system.[32] Our sample includes data from years: 2006, 2007 and 2008. The Current Population Survey (CPS) is sponsored by both the U.S. Census Bureau and the Bureau of Labor Statistics (BLS) (U.S. Census Bureau, 2006). It is a comprehensive household survey, providing a range of data on demographics, labor force, earnings, etc. The CPS takes place each month during the calendar week that contains the 19th of the respective

[31] Available at the following web page: http://www.bls.gov/tus/.
[32] Available at the following web page: http://www.census.gov/cps/data/.

month. The reference week, the week about which the questions are asked, is the week previous to the week when the interview is conducted, hence it is the calendar week that contains the 12th of the respective month (U.S. Census Bureau, 2006). Each household is interviewed eight times. Four interviews take place once per month in a sequence of four months, and four interviews take place the following year during the same four months. While the sample size of the CPS has varied throughout the years because of budgetary and other issues, the monthly sample size is approximately 50,000 households (USDA – ERS). Each household that is surveyed is identified with a fourteen digit code. It is the household code that allows us to match the households from all the three samples into one dataset.

The CPS contains several additional questions as part of a variety of Supplements, amongst which the Food Security Supplement (FSS). The Supplements usually ask questions of public interest and are sponsored by various governmental and non-governmental institutions. The questions of interest related to our study in this survey are: level of weekly food expenditures and household SNAP participation. The FSS was added to the CPS in 1995, and it only takes place in one month (generally in December) during the year.[33] The sample of 44,000 households surveyed in 2008 is representative of the U.S. civilian population of approximately 118 million households (Nord, Andrews and Carlson, 2009).[34] The interview is conducted with an adult member of the family (Nord, Andrews and Carlson, 2009). The main aim of this supplement is to understand issues related to food insecurity in the United States. All the household characteristics' variables included in our model are obtained from the CPS-FSS data. Those include: age, gender, race, level of education, metropolitan status, employment status,

[33] From 1995 to 2001, the Food Security Supplement survey was conducted in different months; however since 2001 it has been continuously conducted in December.
[34] In 2006, the number of households surveyed was 46,500 households (Nord, Andrews and Carlson, 2007).

wage rate, hours worked per week, household income, number of children under 18 years old, household food expenditures and household SNAP participation in the previous 12 months.

The American Time Use Survey (ATUS) is sponsored by the BLS and is conducted on a continuous basis since January 2003, by the U.S. Census Bureau (BLS, 2012). The ATUS is designed and conducted in order to gather information on how people aged 15 and older, living in the United States of America, spend their time. It includes a set of questions regarding the time people spend in activities such as: working, eating, sleeping, watching television, socializing, doing volunteer work, etc. in a regular day. The ATUS respondents are a subset of the respondents who complete the Current Population Survey (BLS, 2010a). The ATUS sample includes people living across the whole United States of America, age 15 and older; but excludes active military personnel and people residing in prisons and nursing homes (BLS, 2010a). While the questions asked pertain both to the individual respondent as well as to his or her respective household, only one respondent per household is interviewed. As in the case of CPS-FSS, the respondent of the ATUS is not necessarily the head of the household. In the case of the ATUS, the respondent can be any member of the household that is 15 years or older. While approximately 7,500 households retire from the CPS each month, about 3,375 of those households got to become part of the ATUS sample in 2003. This totaled to 40,500 households annually (BLS, 2010a). However, beginning December 2003 the ATUS sample was reduced by 35% from a total of 3,375 households per month to only 2,194 households per month. The ATUS survey is conducted using a computer assisted telephone interviewing (CATI) system (BLS, 2010a).

The demographic data for each household that completes the ATUS comes from the CPS, as the respondents of the former are a subset of the respondents of the later. However, such information may be reviewed during the ATUS interview in case any changes have occurred between the final (eighth) month of the CPS and the time when the respondent completes the ATUS (BLS, 2010a). The major contribution of the ATUS is indicating how much time, if any, the respondents spend in different activities in a typical pre-specified day (as explained above). Hence, in order to understand the variables in ATUS datasets, one needs to understand the different types of activities that the respondents report to have been engaged in. All the activities are separated into seventeen subgroups (BLS, 2009). Our variables of interest for the purpose of this study come from the following subgroups: Household Activities, Consumer Purchases, Traveling and Eating and Drinking.

The activities of special interest to our study include: food and drink preparation (020201), food presentation (020202), kitchen and food cleanup (020203), grocery shopping (070101) and travel related to grocery shopping (180701) (BLS, 2004). The sum of the time spent in all these activities makes up to total time spent in food production.[35] We are also interested to analyze the elasticity of substitution when consumption time is included in the analysis. In that case, we also add the time spent in eating and drinking (110101).[36]

As previously indicated, in addition to household level data, we also need state level data on unemployment rates and SNAP payment errors (overpayment rates and underpayment rates). Data on state level unemployment rates during the study period are obtained from BLS – Local Area Unemployment Statistics.[37] Data on SNAP payment error rates during the analysis period

[35] Note that the ATUS respondent is not always the main meal preparer in the household, hence these estimates might be a little bit lower than the actual time spent in household food production. Based on the Eating and Health Module (variable *euprpmel*), 1,393 respondents (i.e. 82.43% of the sample) report to be the main meal preparers in their respective households.

[36] Note that ATUS does not distinguish between time spent consuming meals prepared at home versus time spent consuming meals at restaurants.

[37] Available at the following web page: http://www.bls.gov/lau/#tables.

are obtained from the USDA – FNS. The error rates are reported as percentage (over-paid or under-paid). Because of the large budget dedicated to paying SNAP benefits, even small percentage overpayments and underpayments translate in huge amounts of money – hence the states are continuously given incentives by the Federal government to decrease the payment error rates. For example, in 2010 the overall payment error rate reached a record low of 3.81 (overpayment rate was 3.05 while the underpayment rate was 0.75), which reportedly led to a $356 million reduction in incorrect payments (USDA News Release, 2011). However, as explained in the literature review section, reductions in underpayment and overpayment rates, in addition to affecting the number of participants directly, also affect the number of participants indirectly by making recertification periods shorter and hence increasing the transaction costs of being a participant.

Following Baral, Davis and You (2011) we limit our analysis to single-headed households between 18 and 64 years old, and keep only observations for which time in food production, time in food consumption and weekly food expenses are all greater than zero. The sample size includes 1,872 observations, of which 1,753 observations have non-missing values on SNAP participation. Of those that report whether or not their household received SNAP benefits in the previous 12 months, 226 households (i.e. about 12.9% of the sample) are SNAP participants. This is comparable to the percentage of the U.S. population that receives SNAP benefits. For example, in 2011 approximately 14% of the population received SNAP benefits at some point during the year, while in 2008 approximately 9% of the population did (Author's calculations). Table 2.1 includes summary of statistics for the variables used in the analysis.

Table 2. 1. Summary Statistics

	The Whole Sample			SNAP Participants			SNAP Non-participants		
	Observations	Mean	S.D.	Observations	Mean	S.D.	Observations	Mean	S.D.
Weekly Wage	345	790.17	512.30	27	322.67	185.27	301	832.56	513.75
SNAP	1,753	0.13	0.34	226	--	--	1,527	--	--
MealPrep	1,872	0.82	0.38	226	0.88	0.32	1,527	0.81	0.39
FoodExp (in $)	1,872	113.07	84.63	226	113.63	94.86	1,527	112.98	81.08
FoodPrepTime (in min.)	1,872	64.56	61.72	226	80.50	62.84	1,527	61.81	60.15
EatingTime (in min.)	1,872	123.54	75.38	226	130.74	77.18	1,527	121.56	73.43
NonWhite	1,872	0.23	0.42	226	0.38	0.49	1,527	0.21	0.41
Female	1,872	0.64	0.48	226	0.84	0.37	1,527	0.62	0.49
Employed	1,872	0.75	0.43	226	0.40	0.49	1,527	0.81	0.39
Weekly Work Hours	1,287	40.99	9.82	78	32.58	10.00	1,132	41.50	9.67
HouseholdIncome	1,708	43,401.64	35,906.33	211	12,719.19	10,427.80	1,497	47,726.29	36,115.74
Metropolitan	1,857	0.84	0.37	225	0.78	0.42	1,513	0.84	0.36
HigherEducation	1,872	0.63	0.48	226	0.38	0.49	1,527	0.66	0.47
Age	1,872	45.51	11.38	226	42.17	11.75	1,527	45.89	11.31
NoOfChildren	1,872	0.64	1.03	226	1.31	1.43	1,527	0.56	0.95
SNAPOverpayment	1,872	4.34	1.46	226	4.38	1.41	1,527	4.33	1.46
SNAPUnderpayment	1,872	1.11	0.51	226	1.09	0.50	1,527	1.11	0.52
StateUnemployment	1,872	5.03	1.09	226	5.12	1.10	1,527	5.02	1.09

(*Sources:* CPS, ATUS, FSS, USDA - FNS, and BLS)

SNAP participants compose approximately 13 percent of our sample. While not all respondents report to be the main meal preparer, 88 percent of respondents from households that received SNAP benefits report to be the main meal preparer compared to 81 percent of non-participants. Interestingly, there is no real difference in the amount of weekly food expenditures between SNAP participants and nonparticipants. However, SNAP participants spent more time on average on home food production compared to nonparticipants. While SNAP participants report to spend on average 1 hour and 21 minutes per day engaging in food production at home, for nonparticipants the figure is a approximately 20 minutes less. When both food preparation time and consumption time are combined ("eating time"), the difference between participants and nonparticipants is 9 minutes, while the sample average eating time is a little over 2 hours.

As indicated in the literature, females and non-whites are more likely to be SNAP participants. Thirty-eight percent of those households that report to receive SNAP benefits are not white, while only 23 percent of the sample is not white. As expected, SNAP participants are less likely to be employed, work fewer hours per week and have lower household incomes compared to nonparticipants. While only approximately 40 percent of SNAP participants report to be employed, the figure for nonparticipants is 81 percent. The sample average annual household income is $43,401, but the average for SNAP participants is $12,719.[38] Households in metropolitan regions are less likely to be SNAP participants as 78 percent of SNAP participants are from metropolitan regions, compared to 84 percent of nonparticipants. SNAP participants are also on average less likely to have completed any higher education (beyond high school) and are younger than nonparticipants. Households that report to be SNAP participants also have a higher

[38] Note that we use a categorical variable on household income (FSS_hufaminc) to create a continuous variable based on the median value of income for each category. Hence, the reported values are approximations of household income rather than actual reported values.

number of children under 18 years old, compared to nonparticipants. Finally, the difference on the average rates of SNAP overpayment, SNAP underpayment and state unemployment levels are relatively small between the two subgroups. However, they have the expected direction (i.e. SNAP overpayment rate is higher, SNAP underpayment rate is lower and state unemployment rate is higher for the SNAP participants subgroup).

2.5. Empirical Results

The empirical analysis was done in two ways. The first way involves estimating the elasticity of substitution for each group (SNAP participants and nonparticipants) by limiting the analysis sample to that specific group. By splitting the sample according to SNAP participation status, we avoid the selection bias. The second way involves making use of the propensity score matching method in order to adjust for endogeneity and estimate the impact of treatment (SNAP participation). The results of the first step are reported in Table 2.2 while the results of the second step are reported in Table 2.4. Note that all results are reported for both cases when consumption time is included in as well as excluded from the analysis.

Table 2.2 includes four different estimation models. Model 1 presents the estimation result for the benchmark sample, only excluding observations for which there are missing data on SNAP participation.[39] Model 2 contains the results for the SNAP participants' subsample, while Model 3 contains the results for the SNAP nonparticipants' subsample. Finally, as indicated in the previous section, not all respondents report to be the main meal preparers in their respective households. In the benchmark sample, 1,536 respondents (82.05%) report to be the

[39] The sample reduces to 1,753 observations since 119 observations have missing data on SNAP participation.

main meal preparer, while 336 (17.95%) do not report to be the main meal preparer.[40] Hence, in Model 4 of Table 2.2. we report the elasticity of substitution for main meal preparers versus non main meal preparers.

[40] Of the 336 who do not report to be the main meal preparer in their households, 233 are not the main meal preparers, 85 split this duty equally with other household members, 14 refused to answer the question, and 1 reported to not know the answer.

Table 2. 2. Goods-Time Elasticity of Substitution Results Before Propensity Score Matching

	Without Consumption Time				With Consumption Time			
	Model 1	Model 2	Model 3	Model 4	Model 1	Model 2	Model 3	Model 4
LnWage	0.475***	-0.146	0.433***	0.579***	0.282***	-0.118	0.290***	0.457***
	(0.107)	(0.318)	(0.120)	(0.202)	(0.091)	(0.229)	(0.102)	(0.149)
*LnWage*MealPrep*				-0.371				-0.433**
				(0.226)				(0.178)
MealPrep				3.81				4.770**
				(2.793)				(2.182)
Children	0.191***	-0.076	0.294***	0.105***	0.169***	-0.000	0.227***	0.108***
	(0.041)	(0.060)	(0.047)	(0.037)	(0.028)	(0.050)	(0.031)	(0.027)
Constant	-4.983***	2.418	-4.494***	-5.710**	-3.563***	1.298	-3.680***	-5.300***
	(1.309)	(3.786)	(1.478)	(2.499)	(1.118)	(2.744)	(1.252)	(1.825)
Observations	1,753	226	1,527	1,872	1,753	226	1,527	1,753
		(SNAP participants)	(SNAP nonparticipants)			(SNAP participants)	(SNAP nonparticipants)	
R-squared	0.042	0.01	0.063	0.107	0.042	0.001	0.06	0.105

Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1

The elasticity estimation results reported in Model 1 are similar to those obtained by Baral, Davis and You (2011). Namely, the estimated elasticity of substitution when consumption time is excluded is $\sigma_f = 0.48$, while when consumption time is included is $\sigma_e = 0.28$.[41] Both results are statistically significant at 1 percent level of significance.

The benchmark sample includes 226 observations of SNAP participants. In Model 2, by limiting the sample to only SNAP participants and carrying on the same empirical analysis, we obtain the goods-time elasticity of substitution for SNAP participants. The estimated coefficient has a negative sign implying that these two factors are complements; however the high standard error indicates that the result is not statistically significant.

We conduct the same analysis for SNAP nonparticipants and report the results in Model 3. For SNAP nonparticipants, the estimated elasticity of substitution when consumption time is not included in the analysis is $\sigma_f = 0.43$ while when consumption time is included in the analysis is $\sigma_e = 0.29$. These findings are consistent with the findings by Baral, Davis and You (2011) since they indicate that the elasticity of substitution is higher when only food production time is considered compared to when the time on the eating occasion is considered. Note however that the estimated elasticity of substitution is lower for SNAP nonparticipants than for the full sample in the case when consumption time is excluded from the analysis and slightly higher when consumption time is included in the analysis.

When considering the variables on time spent in household food production and food consumption, there are two data issues that are worth discussing. The first issue is that consumption time does not include only the time spent consuming food prepared at home.

[41] We use the same notation for the elasticity of substitution as in Baral, Davis and You (2011).

Hence, potentially, part of the time spent in food consumption is spent consuming meals prepared away from home (i.e. in restaurants). However, this is an issue that we will ignore in our study since unfortunately the ATUS currently does not ask respondents to clarify whether the food consumed during the indicated time was prepared at home or away from home.[42] However, we still report the analysis for the case when consumption time is included for comparison as well as to keep a similar structure to the study conducted by Baral, Davis and You (2011).

The second data issue of interest is that survey respondents are neither necessarily the head of the household, nor necessarily the main meal preparer in the household. As indicated in the previous section, an ATUS respondent may be any member of the household who is 15 years or older. While our sample is limited to single-headed households where the respondent reports to be between 18 and 64 years old, there are still some cases when the respondent does not report to be the main meal preparer.[43] Model 4 of Table 2.2. reports the results of the estimated elasticity of substitution for main meal preparers versus non main meal preparers. The elasticity of substitution is lower for the main meal preparers compared to non main meal preparers in both cases (when consumption time is included as well as when consumption time is excluded from the analysis). However, in the case when consumption time is not included in the analysis - which is the main case we're concerned with in this study - the difference is not statistically significant. While it would be interesting to estimate the elasticity of substitution for SNAP participants and nonparticipants when the sample is limited to only main meal preparers, that would lead to a significant reduction in our sample size. Hence for these reasons, we will not

[42] Note that the Eating and Health module asks further questions about time spent in eating and drinking, but it does not specify where such food was prepared. The goal of such questions is to understand how much time respondents spent engaging in primary eating and drinking versus secondary eating and drinking (i.e. eating and drinking while watching TV, etc.).

[43] The Eating and Health Module allows us to identify the main meal preparers versus non main meal preparers.

limit the sample size to the main meal preparers, but it is still interesting to note that the estimates might be somewhat different for this subgroup.

As previously indicated, the second way we conduct the empirical analysis is to make use of the propensity score matching method in order to adjust for self-selection into SNAP program and create counterfactual comparison. Note that the sample size is further limited to only include observations with non-missing values of the covariates used in the propensity score matching analysis. This leads to a sample size of 1,693 observations. Table 2.3 below reports the summary statistics of the covariates used in the model specification of propensity score matching. This combination of covariates fulfilled the balance requirement of propensity score matching. The balancing requirement test checks if observations which have the same propensity score also have the same distribution of the independent variables regardless of whether the observations belong to the treatment group or to the control group (Lee, 2006). The first two columns present the summary of statistics for the treatment group (i.e., SNAP participants). The third and fourth columns present the summary of statistics for the control group (i.e., SNAP nonparticipants). And the fifth column presents the normalized differences between the covariate distributions between the treatment and control groups, which were calculated following Abadie and Imbens (2011).

Table 2. 3. Summary Statistics After Propensity Score Matching

	Treated (210)		Control (1,483)		T-Test Results (Means' Comparison)	Normalized Difference
	Mean	SD	Mean	SD		
List of Covariates						
NonWhite	0.367	0.483	0.211	0.408	-1.41	0.246
Female	0.838	0.369	0.616	0.487	-3.35	0.364
Employed	0.395	0.490	0.813	0.390	16.00	-0.666
Metropolitan	0.781	0.415	0.842	0.365	3.11	-0.111
HigherEducation	0.386	0.488	0.670	0.470	5.86	-0.420
NoOfChildren	1.338	1.446	0.557	0.934	-8.33	0.454
Age	41.805	11.717	45.931	11.332	3.12	-0.253
HouseholdIncome	12,738	10,449	47,857	36,204	16.29	-0.932
SNAPOverpayment	4.367	1.440	4.327	1.451	-1.87	0.019
SNAPUnderpayment	1.080	0.505	1.114	0.517	0.07	-0.047
Outcomes:						
Ln(FoodExp/FoodPrepTime)	0.437	1.243	0.791	1.294		-0.198
Ln(FoodExp/EatingTime)	-0.239	0.987	-0.138	1.002		-0.072

The difference in the distribution of the covariates between the treatment and control groups is large for variables such as employment status and household income. Other covariates such as SNAP overpayment and underpayment rates, as well as whether the respondent lives in a metropolitan or nonmetropolitan area have a smaller difference in the distribution between the two groups. According to Abadie and Imbens (2011), the normalized differences between covariates for the two groups provide a useful measure of the difficulty in adjusting for differences in the covariates used in the propensity score matching method. Covariates for which the difference in the distribution is larger affect the balancing requirement of the propensity score, as briefly mentioned above.

The results on the goods-time elasticity of substitution after applying the propensity score matching weighting to the sample was applied are reported in Table 2.4. Results from models with and without consumption time are reported. In order to access the robustness of our results, we applied five different methods of propensity score matching, namely: One-to-one matching, K-nearest neighbor matching, Radius matching, Kernel matching and Mahalanobis matching. The different ways to conduct propensity score matching differ in the process by which the matching between the participants and nonparticipants is done. For example, the nearest neighbor matching randomly orders all observations, selects a treated observation (participant) and matches it with a non-treated observation (nonparticipant) with the closest propensity score. Radius matching uses the nearest neighbor within each specified caliper and all the observations in the non-treatment group within each caliper. Kernel matching method matches each observation in the treatment group with a weighted average of the observations with similar propensity score in the control group, where greater weight is assigned to observations with a closer propensity score. Finally, the Mahalanobis matching randomly orders all observations, calculates the so-called “Mahalanobis distance” between the first treated observation and all non-treated observations, chooses the non-treated observation with the shortest distance from the first treated observation and then removes them both from the pool, repeating the process again. Since each of the matching methods has its advantages and disadvantages, we employ all five for robustness check. For all the matching methods we specify the caliper value at two thresholds, 0.01 and 0.001. The caliper option specifies the maximum distance of controls. All the five propensity score methods considered provide the exact same coefficient estimations, as reported in Case 1 of Table 2.4.

In addition to the caliper specification, an additional option in the propensity score methodology is limiting the analysis to the observations on common support. Treatment observations are on common support if their respective propensity scores are lower or equal to the maximum and higher than or equal to the minimum propensity score for the controls. A more restrictive caliper specification (i.e. caliper of 0.001) leads to a higher number of observations to be out of common support. Hence, limiting the analysis to only the subsample of observations on common support leads to slightly different results and a lower sample size. Such results vary across different propensity score methods, but we report the results obtained from Radius matching in Table 2.4. Case 2 reports the results when caliper specification is 0.01 and Case 3 when caliper specification is 0.001. Figures 2.1 and 2.2 below show the region of common support for the Radius matching for caliper 0.01 and 0.001 respectively. As expected, the region of common support is smaller as the caliper (distance) specification becomes more restrictive.

Figure 2. 1. Radius Matching (Caliper 0.01)

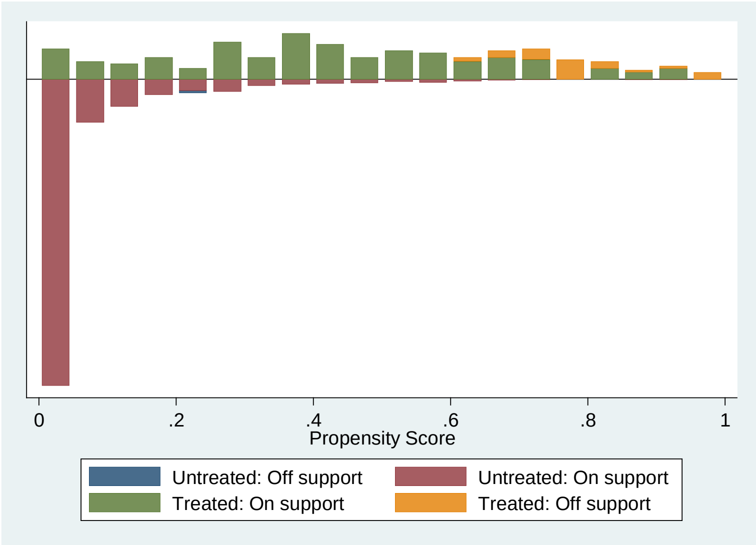


Figure 2. 2. Radius Matching (Caliper 0.001)

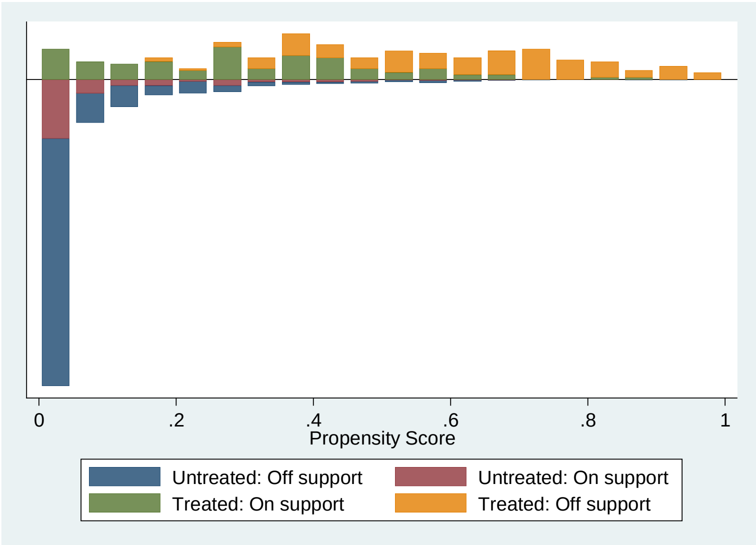


Table 2. 4. Goods-Time Elasticity of Substitution Results After Propensity Score Matching

	Without Consumption Time			With Consumption Time		
	Case 1	Case 2	Case 3	Case 1	Case 2	Case 3
LnWage	0.861*	0.899*	0.453	0.373	0.413	0.185
	(0.505)	(0.502)	(0.680)	(0.404)	(0.392)	(0.502)
*LnWage*SNAP*	-0.987	-0.931	-1.034	-0.537	-0.497	-0.594
	(0.604)	(0.605)	(0.815)	(0.468)	(0.447)	(0.604)
*LnWage+LnWage*SNAP*	-0.126	-0.032	-0.582	-0.164	-0.084	-0.409
	(0.337)	(0.380)	(0.511)	(0.242)	(0.271)	(0.370)
SNAP	12.103*	11.429	12.614	6.818	6.339	7.414
	(7.332)	(7.306)	(9.857)	(5.713)	(5.443)	(7.314)
Children	-0.065	-0.055	0.142	-0.013	-0.001	0.104
	(0.149)	(0.174)	(0.103)	(0.116)	(0.134)	(0.093)
Constant	-9.943	-10.401*	-5.028	-4.949	-5.431	-2.613
	(6.093)	(6.055)	(8.162)	(4.880)	(4.726)	(6.036)
Observations	1,693	1,656	460	1,693	1,656	460
R-squared	0.042	0.042	0.074	0.024	0.023	0.055

*Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1*

After applying the propensity score matching to adjust for endogeneity, the goods-time elasticity results vary from the results obtained before the propensity score matching, but tell the same story. When the common support condition is not required in the regression analysis (Case 1), the estimated elasticity of substitution for SNAP nonparticipants is $\sigma_f = 0.86$ (p-value 0.09) when consumption time is not included, and $\sigma_e = 0.37$ (p-value 0.36) when consumption time is included. While these results are different from those obtained by just limiting the sample to SNAP nonparticipants as reported in Table 2.2 (Case 3), they are consistent with those results in that they indicate that the elasticity of substitution is higher when consumption time is not included versus when consumption time is included.

The elasticity of substitution for SNAP participants is not statistically different from zero. This result holds across all the cases considered, both with consumption time and without consumption time. In order to increase the confidence in the statistical inference results, we apply the jackknife method to derive robust estimates of the standard errors. This result was somewhat expected since when the analysis was conducted by limiting the sample to SNAP participants only, the results showed that the elasticity of substitution for this subgroup is not statistically different from zero. This result has important policy implications as it indicates that the SNAP participants face Leontief production function in household food production.

2.6. Conclusions

This study finds that the goods-time elasticity of substitution in household food production for SNAP participants is not statistically different from zero. This indicates that SNAP participants need a fixed amount of both inputs to produce food at home; hence increasing the amount of money alone will not lead to more food production if time allocated to home food production remains unchanged. This result has important policy implications for policies related to SNAP. A zero elasticity of substitution between the inputs of home food production implies that policy changes that lead to an increase of benefits paid to SNAP participants will have limited effectiveness if such households do not also simultaneously increase the time dedicated to home food production. In other words, simply pouring more money into the program is only a necessary, but not sufficient, condition to increase household food production for SNAP participants. The other necessary condition is for the households to also increase the time dedicated to home food production, which will come at a cost that households have to incur.

As previously indicated, SNAP is the most important food assistance program with millions of households nationwide receiving benefits from it. For a program of this magnitude, the implications of our results affect a significant portion of the population – hence it is important for researchers to continue exploring this topic. An extension to this study may be to estimate the elasticity of substitution for SNAP participants by limiting the sample to main meal preparers only. Another extension is to extend the analysis to a broader sample by not restricting it to single headed households as in this case or by extending the period of analysis. It would also be interesting to estimate the elasticity of substitution pre- and post- the implementation of the ARRA provisions which increased the amount of benefits paid to SNAP participants in the last few years during the economic crisis.

References

Abadie, A. and G. W. Imbens. 2011. "Bias-Corrected Matching Estimators for Average Treatment Effects." *Journal of Economics & Business Statistics* 29(1).

Baral, R., G. C. Davis, and W. You. 2011. "Consumption time in household production: Implications for the goods-time elasticity of substitution." *Economics Letters*, doi:10.1016/j.econlet.2011.03.005.

Becker, G. S. 1965. "A Theory of the Allocation of Time." *The Economic Journal* 75(299): 493-517.

Blackorby, C, and R. R. Russell. 1981. "The Morishima Elasticity of Substitution: Symmetry, Constancy, Separability, and its Relationship to the Hicks and Allen Elasticities." *The Review of Economic Studies* 48(1): 147-158.

_______. 1989. "Will the Real Elasticity of Substitution Please Stand Up? (A Comparison of the Allen/Uzawa and Morishima Elasticities)." *The American Economic Review* 79(4): 882-888.

Bureau of Labor Statistics. 2004. "American Time Use Survey Lexicon." Retrieved 2011 from http://www.bls.gov/tus/lexiconnoex2004.pdf.

_______. 2009. "American Time Use Survey (ATUS) Coding Rules." Retrieved 2011 from http://www.bls.gov/tus/tu2009coderules.pdf.

_______. 2010a. "American Time Use Survey User's Guide. Understanding ATUS 2003 to 2009." Retrieved 2011 from http://www.bls.gov/tus/atususersguide.pdf.

_______. 2010b. "Eating and Health Module Questionnaire." Retrieved 2011 from http://www.bls.gov/tus/ehmquestionnaire.pdf.

_______. 2012. "American Time Use Survey – Frequently Asked Questions." Retrieved 2011 from http://www.bls.gov/tus/atusfaqs.htm.

_______. "Labor Force Statistics from the Current Population Survey." Retrieved 2011 from http://www.bls.gov/cps/home.htm.

_______. "Local Area Unemployment Statistics. Retrieved 2012 from http://www.bls.gov/lau/#tables.

Caliendo, M. and S. Kopeinig. 2005. "Some Practical Guidance for the Implementation of Propensity Score Matching." IZA DP No. 1588. Retrieved 2012 from http://ftp.iza.org/dp1588.pdf.

Carlson, A., M. Lino, W. Juan, K. Hanson, and P. P. Basiotis. 2007. "Thrifty Food Plan, 2006." Retrieved 2012 from http://www.cnpp.usda.gov/Publications/FoodPlans/MiscPubs/TFP2006Report.pdf.

Center on Budget and Policy Priorities. 2012. "Policy Basics: Introduction to the Supplemental Nutrition Assistance Program (SNAP)." Retrieved 2012 from http://www.cbpp.org/cms/index.cfm?fa=view&id=2226.

Currie, J., J. Grogger, G. Burtless and R. F. Schoeni. 2001. "Explaining Recent Declines in Food Stamp Program Participation [with Comments]." *Brookings-Wharton Papers on Urban Affairs*: 203 – 244.

Davis, G. C., and C. R. Shumway. 1996. "To tell the Truth about Interpreting the Morishima Elasticity of Substitution." *Canadian Journal of Agricultural Economics* 44: 173-182.

Davis, G. C., and W. You. 2010a. "The Thrifty Food Plan In Not Thrifty When Labor Cost Is Considered." *The Journal of Nutrition Issues and Opinions*, doi: 10.3945/jn.109.119594.

_______. 2010b. "The time cost of food at home: general and food stamp participant profiles." *Applied Economics* 42: 2537-2552.

Diewert, W.E. 1974. "A note on the Aggregation and Elasticities of Substitution." *The Canadian Journal of Economics* 7(1): 12-20.

Eating and Health Module User's Guide. 2010. Retrieved 2011 from http://www.ers.usda.gov/Publications/AP/AP047/AP047.pdf.

Hamermesh, D. S. 2008. "Direct estimates of household production." *Economics Letters* 98:31-34.

Kabbani, N. S., and P. E. Wilde. 2003. "Short Recertification Periods in the U.S. Food Stamp Program." *The Journal of Human Resources* 38: 1112-1138.

Kornfeld, R. 2002. "Explaining Recent Trends in Food Stamp Program Caseloads." E-FAN No.(02-008). Retrieved 2012 from http://www.ers.usda.gov/publications/efan02008/.

Laskawy, T. 2011. "Stamping out hunger with fast food?" Retrieved 2012 from http://grist.org/food/2011-09-08-stamping-out-hunger/.

Lee, Wang-Sheng. 2006. "Propensity Score Matching and Variations on the Balancing Test." Retrieved 2012 from https://editorialexpress.com/cgi-bin/conference/download.cgi?db_name=esam06&paper_id=217.

Mykerezi, E. and B. Mills. 2010. "The Impact of Food Stamp Program Participation on Household Food Insecurity." *American Journal of Agricultural Economics* 92(5): 1379-1391.

Nichols, A. 2007. "Causal inference with observational data." *The Stata Journal* 7(4): 507-541.

_______. 2008. "Erratum and discussion of propensity-score reweighting." *The Stata Journal* 8(4): 532-539.

Nord, M., M. Andrews, and S. Carlson. 2007. "Household Food Security in the United States, 2006." Retrieved 2012 from http://www.ers.usda.gov/Publications/ERR49/ERR49.pdf.

_______. 2009. “Household Food Security in the United States, 2008.” Retrieved 2012 from http://www.ers.usda.gov/Publications/ERR83/ERR83.pdf.

Rosenbaum, P. R. and D. B. Rubin. 1983. “The Central Role of the Propensity Score in Observational Studies for Causal Effects.” *Biometrica* 70 (1): 41-55.

Silberberg, E. 1978. “The Structure of Economics.” McGRAW-HILL, INC.

United States – Census Bureau. 2006. “Current Population Survey. Design and Methodology – Technical Paper 66.” Retrieved 2011 from http://www.census.gov/prod/2006pubs/tp-66.pdf.

_______. 2008. “Current Population Survey. Food Security Supplement File – Technical Documentation, CPS -08.” Retrieved 2012 from http://www.census.gov/apsd/techdoc/cps/cpsdec08.pdf.

USDA – Economic Research Service. Eating and Health Module (ATUS): Overview. Retrieved 2011 from http://www.ers.usda.gov/Data/ATUS/Overview.htm.

USDA – Food and Nutrition Service. 2012a. “A short history of SNAP.” Retrieved 2012 from http://www.fns.usda.gov/snap/rules/Legislation/about.htm.

_______. 2012b. “American Recovery & Reinvestment Act of 2009 – Supplemental Nutrition Assistance Program.” Retrieved 2012 from http://www.fns.usda.gov/fns/recovery/recovery-snap.htm.

_______. 2012c. “Program Data – Supplemental Nutrition Assistance Program.” Retrieved 2012 from http://www.fns.usda.gov/pd/snapmain.htm.

_______. 2012d. “SNAP Eligibility.” Retrieved 2012 from http://www.fns.usda.gov/snap/applicant_recipients/eligibility.htm.

_______. 2012e. “SNAP Quality Control.” Retrieved 2012 from http://www.fns.usda.gov/snap/qc/default.htm.

USDA News Release. 2011. “Nation’s Primary Food Assistance Program Reaches Highest Accuracy Rate in History of the Program.” Release No. 0256.11.

Wooldridge, J.M. 2009. "Introductory Econometrics: A Modern Approach." Mason, OH: South-Western Cengage Learning.

Wooldridge, J.M. 2010. “Econometric Analysis of Cross Section and Panel Data.” 2nd edition. MIT Press.